

RL Vaccine 11:438-444(1993).
 CC -!- FUNCTION: This protein directs fusion of viral and cellular
 CC membranes.
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
 CC LINKED BY A DISULFIDE BOND.
 CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M21849; AAA42878.1; -
 CC DR EMBL; X55509; CAA46481.1; -
 CC DR PIR; J50321; VGNZCD.
 CC DR PIR; S21382; S21382.
 CC DR HSSP; P04849; LSVP.
 CC DR InterPro; IPR000776; Fusion gly.
 CC DR Pfam; PF00523; fusion_gly; 1.
 CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
 CC FT SIGNAL 1 7
 CC FT CHAIN ? 662 FUSION GLYCOPROTEIN F0.
 CC FT CHAIN ? 224 PROTEIN F2.
 CC FT CHAIN 225 662
 CC FT TRANSMEM 606 629
 CC FT DISULFID 180 307 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
 CC FT CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC FT CONFLICT 3 3 R -> K (IN REF. 2).
 CC FT CONFLICT 140 140 D -> N (IN REF. 2).
 CC FT CONFLICT 152 152 N -> S (IN REF. 2).
 CC FT CONFLICT 171 171 I -> M (IN REF. 2).
 CC FT CONFLICT 174 174 A -> V (IN REF. 2).
 CC FT CONFLICT 662 662 L -> H (IN REF. 2).
 CC SQ SEQUENCE 662 AA; 72970 MW; F82C81C9797805F0 CRC64;
 Query Match 37.9%; Score 66; DB 1; Length 662;
 Best Local Similarity 73.7%; Pred. No. 0.14;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 15 SLKLLSEIKGVVHRLEGV 33
 Db 396 SYPTLSEVKGIVVHRLEAV 414
 RESULT 8
 ID_GON1_SPAU STANDARD; PRT; 95 AA.
 AC P51979;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I)
 DE (LH-RH I) (Luliberin I) (SBNRH).
 GN GNRH1.
 OS Sparus aurata (Gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Percoidae;
 OC Sparidae; Sparus.
 CC NCBI_TaxID=8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95268499; PubMed=7749463;
 RA Gethif Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
 RT "Molecular cloning and characterization of a novel gonadotropin-

RT releasing hormone from the gilthead seabream (Sparus aurata).";
 RL Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
 RN [2]
 RP SEQUENCE OF 26-35.
 RC TISSUE=Brain;
 RX MEDLINE=95083645; PubMed=7991588;
 RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
 RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
 RT "Three forms of gonadotropin-releasing hormone characterized from
 RT brains of one species.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MASS SPECTROMETRY: MW=1113.6; METHOD=MAJDI; RANGE=26-35.
 CC -!- SIMILARITY: Belongs to the GNRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U30320; AAA75469.1; -
 CC DR InterPro; IPR002012; GNRH.
 CC DR InterPro; IPR004079; GonadoliberinI.
 CC DR Pfam; PF00446; GNRH; 1.
 CC DR PRINTS; PRO1541; GONADOLIBRNI.
 CC DR PROSITE; PS00473; GNRH; 1.
 CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Signal; Multigene family; Pyrrolidone carboxylic acid.
 CC FT SIGNAL 1 25
 CC FT CHAIN 26 95 PROGONADOLIBERIN I.
 CC FT PEPTIDE 26 35 GONADOLIBERIN I.
 CC FT PEPTIDE 39 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
 CC FT MOD RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
 CC SQ SEQUENCE 95 AA; 10753 MW; 49313FD6FD68B7DA CRC64;
 Query Match 37.6%; Score 65.5; DB 1; Length 95;
 Best Local Similarity 45.5%; Pred. No. 0.019;
 Matches 15; Conservative 3; Mismatches 14; Indels 1; Gaps 1;
 QY 2 HWSYGLRPGSSGPSLKLSEIKGVVHRLEGE 34
 Db 27 HWSYGLSPGKG-RDLDSLSDTLGNIIERPHVD 58
 RESULT 9
 ID_VGLF_RINDK STANDARD; PRT; 546 AA.
 AC P12574;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
 DE Fusion glycoprotein F1].
 GN F.
 OS Rinderpest virus (strain Kabete O) (RDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
 CC NCBI_TaxID=11242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88322864; PubMed=3413983;
 RA Hsu D., Yamanaka M., Miller J., Dale B., Grubman M., Yilma T.;
 RT "Cloning of the fusion gene of rinderpest virus: comparative sequence
 RT analysis with other morbilliviruses.";
 RL Virology 166:149-153(1988).
 CC -!- FUNCTION: This protein directs fusion of viral and cellular
 CC membranes.
 CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2

```

CC CC LINKED BY A DISULFIDE BOND.
CC CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC CC family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; M21514; AAA47400.1; --
CC CC PIR; A31051; VGNZK.
CC CC HSSP; P04849; 1SVF.
CC CC InterPro; IPR000776; Fusion gly.
CC CC Pfam; PF00523; fusion_gly; 1.
CC CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
CC CC FT SIGNAL 1 19
CC CC FT CHAIN 20 546 FUSION GLYCOPROTEIN F0.
CC CC FT CHAIN 20 108 F2 PROTEIN.
CC CC FT CHAIN 109 546 F1 PROTEIN.
CC CC FT DOMAIN 104 108 ARG-RICH (BASIC).
CC CC FT TRANSMEM 109 133 POTENTIAL.
CC CC FT DOMAIN 484 513 ARG/LYS-RICH (BASIC).
CC CC FT DISULFID 64 191 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC CC FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC FT CARBOHYD 57 57 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC SEQUENCE 546 AA; 58662 MW; 476D74DC018CFCF CRC64;
CC CC
CC CC Query Match 37.4%; Score 65; DB 1; Length 546;
CC CC Best Local Similarity 86.7%; Pred. No. 0.16;
CC CC Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC CC
QY 19 LSEIKGVIVHLEGV 33
DB 284 LSEIKGVIVHLEGV 298
-----
RESULT 10
VGLF PHODV
ID VGLF PHODV STANDARD; PRT; 631 AA.
AC P28886;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Phocine distemper virus (PDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate DK88-4A;
RX MEDLINE=92113538; PubMed=1765768;
RA Koevamees J., Blixenkron-Moeller M., Sharma B., Oervell C.,
RA Norrby E.;
RT "The nucleotide sequence and deduced amino acid composition of the
RT haemagglutinin and fusion proteins of the morbillivirus phocid
RT distemper virus."
RL J. Gen. Virol. 72:2959-2966 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Uster/88;
RX MEDLINE=92398437; PubMed=1524494;
RA Curran M.D., Lu Y.J., Rima B.K.;
RA "The fusion protein gene of phocine distemper virus: nucleotide and
RT deduced amino acid sequences and a comparison of morbillivirus fusion

```

```

RT proteins."
RL Arch. Virol. 126:159-169 (1992).
RN [3]
RP SEQUENCE OF 95-631 FROM N.A.
RC STRAIN=Uster/88;
RX MEDLINE=91089508; PubMed=2264246;
RA Curran M.D., Loan D.O., Rima B.K., Kennedy S.;
RA "Nucleotide sequence analysis of phocine distemper virus reveals its
RT distinctness from canine distemper virus."
RL Rec. Rec. 127:430-431 (1990).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES F-1 AND F-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; D10371; BAA01206.1; --
CC CC PIR; A48346; A48346.
CC CC PIR; J01368; VGNZPD.
CC CC HSSP; P04849; 1SVF.
CC CC InterPro; IPR000776; Fusion gly.
CC CC Pfam; PF00523; fusion_gly; 1.
CC CC KW Glycoprotein; Fusion protein; Transmembrane; Envelope protein; Signal.
CC CC FT SIGNAL 1 631
CC CC FT CHAIN ? 631 FUSION GLYCOPROTEIN F0.
CC CC FT CHAIN ? 188 F2 PROTEIN.
CC CC FT CHAIN 194 631 F1 PROTEIN.
CC CC FT DISULFID 149 276 LINKAGE BETWEEN F2 & F1 (POTENTIAL).
CC CC FT TRANSMEM 89 106 POTENTIAL.
CC CC FT TRANSMEM 194 212 POTENTIAL.
CC CC FT TRANSMEM 575 595 POTENTIAL.
CC CC FT CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC FT CONFLICT 63 I -> V (IN REF. 2).
CC CC SEQUENCE 631 AA; 68873 MW; D1FC87CDD426E9B8 CRC64;
CC CC
CC CC Query Match 37.4%; Score 65; DB 1; Length 631;
CC CC Best Local Similarity 68.4%; Pred. No. 0.19;
CC CC Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
CC CC
QY 15 SLKLSEIKGVIVHLEGV 33
DB 365 SYPTLSEVKGVVVHLEAV 383
-----
RESULT 11
GON1 MACMU
ID GON1 MACMU STANDARD; PRT; 67 AA.
AC P55217;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (uteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE-Hypothalamus;
RX MEDLINE=9524501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female rhesus
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S75918; AAB33096.1; -
CC PIR; I78541; I78541; -
CC InterPro; IPR002012; GnRH.
CC InterPro; IPR004079; GonadoliberinL.
CC Pfam; PF00446; GnRH; 1.
CC PRINTS; PR01541; GONADOLIBERNI.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal; Pyrrolidone carboxylic acid.
CC NON_TER 1 1
CC SIGNAL <1 5 BY SIMILARITY.
CC CHAIN 6 >67 PROGONADOLIBERIN I.
CC PEPTIDE 6 15 GONADOLIBERIN I.
CC PEPTIDE 19 >67 GnRH-ASSOCIATED PEPTIDE I.
CC ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY (BY SIMILARITY).
CC MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
CC SIMILARITY).
CC MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
CC SIMILARITY).
CC NON_TER 67 67
CC SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
Query Match 36.8%; Score 64; DB 1; Length 67;
Best Local Similarity 40.0%; Pred. No. 0.022;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 2 HWSYGLRPGSGSPSLKLLSEIKGVI 26
Db 7 HWSYGLRPGGKRDANLMDSFQEIF 31
RESULT 12
ID GONI_HUMAN STANDARD; PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (luliberin I) (Gonadorelin); GnRH-associated
DE peptide I).
DE GN GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366682; PubMed=2671939;

```

```

RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene.";
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:179-183(1996).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT SER-16.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone.";
RL Nature 311:666-668(1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE=83126573; PubMed=6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta.";
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN [5]
RP VARIANT SER-16.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238(1999).
RN [6]
RP EREATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipschutz R., Daley G.Q.,
RA Lander E.S.;
RT "FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
RT the secretion of both luteinizing and follicle-stimulating
RT hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
CC Lutrepulse or Lutrelle (Ferring Pharmaceuticals) and Relisorm
CC (Serono).
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X01059; CAA25526.1; -
CC EMBL; M12578; AAA35916.1; -
CC EMBL; X15215; CAA33285.1; -
CC PIR; S05308; RHUG.
CC Genew; HGNC:4419; GNRH1.
CC MIM; 152760; -
CC GO; GO:0005625; C:soluble fraction; TAS.
CC GO; GO:0005383; F:luteinizing hormone-releasing factor activity; TAS.
CC GO; GO:0007267; F:cell-cell signaling; TAS.
CC GO; GO:0007275; P:development; TAS.
CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR002012; GnRH.

```

```

InterPro: IPR004079; GonadoliberinI.
PFam: PF00446; GnRH: 1.
PRINTS: PR01541; GONADOLIBRN1.
PROSITE: PS00473; GnRH; 1.
Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Placenta; Pharmaceutical; Signal; Polymorphism;
Pyroglutamate carboxylic acid.
SYNOPSIS
  1 23
CHAIN 24 92 PROGONADOLIBERIN I.
  24 92 GONADOLIBERIN I.
PEPTIDE 24 33 GnRH-ASSOCIATED PEPTIDE I.
  37 92 GnRH-ASSOCIATED PEPTIDE I.
PEPTIDE 37 92 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
  26 26 ACTIVITY.
MOD RES 24 24 PYROGLUTAMATE CARBOXYLIC ACID.
MOD RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
VARIANT 16 16 W -> S (in dbSNP:6185).
/FTid=VAR_013943.
SEQUENCE 92 AA; 10380 MW; 30A72221E076FA79 CRC64;
Query Match 36.8%; Score 64; DB 1; Length 92;
Best Local Similarity 40.0%; Pred. No. 0.03;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 2 HWSYGLRPGSGPSLKLISIKGVI 26
  ||||| : : : :
DB 25 HWSYGLRPGGKRDRENLIDSFQEV 49
  ||||| : : : :
RESULT 13
ID VGLF MEASI STANDARD; PRT; 529 AA.
AC P26031; Q83298;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN F.
OS Measles virus (strain IP-3-Ca) (Subacute sclerosing panencephalitis
OS virus).
OC Viruses; ssRNA negative-strand viruses; Nonnegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
NCBI TaxID=11237;
(1)
SEQUENCE FROM N.A.
MEDLINE=922263801; PubMed=1585658;
Schmid A., Spielhofer P., Cattaneo R., Baczko K., Ter Meulen V.,
Billeter M.A.,
RA "Subacute sclerosing panencephalitis is typically characterized by
RT alterations in the fusion protein cytoplasmic domain of the
RT persisting measles virus."
RT Virology 198;910-915(1992).
CC -!- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -!- SUBUNIT: THE MATURE FORM IS A DIMER OF POLYPEPTIDES P-1 AND P-2
CC LINKED BY A DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16566; CAA34567.1; -
CC EMBL; X16566; CAA34568.1; ALT_INIT.
CC HSSP; P04849; ISVP.
CC InterPro: IPR000776; Fusion gly.
CC PFam; PF00523; fusion gly; 1.
CC Glycoprotein; Fusion Protein; Transmembrane; Envelope protein; Signal.

```



```

QY      2 HWSYGLRPGSSGSPSLK-----LSEIKGVIHRLEG 32
      ||||| : |
      ||||| : |
      ||||| : |

DB      25 HWSYGLRPGGKRDRFESLQDWWYHETNEVALPELERLECSVPOSTNVLFGALMNLWLEG 83
      ||||| : |
      ||||| : |
      ||||| : |

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 12.7763 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKIAKMEKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------------|
| 1 | 134.5 | 79.6 | 33 | 1 | US-08-446-692-27 |
| 2 | 134.5 | 79.6 | 33 | 2 | US-08-488-351A-27 |
| 3 | 89 | 52.7 | 21 | 1 | US-08-186-266-6 |
| 4 | 89 | 52.7 | 21 | 1 | US-08-446-692-48 |
| 5 | 89 | 52.7 | 21 | 2 | US-08-488-351A-48 |
| 6 | 89 | 52.7 | 21 | 3 | US-09-100-409A-54 |
| 7 | 89 | 52.7 | 21 | 4 | US-08-464-496-17 |
| 8 | 89 | 52.7 | 21 | 4 | US-08-788-822A-12 |
| 9 | 89 | 52.7 | 21 | 4 | US-08-197-484-97 |
| 10 | 89 | 52.7 | 21 | 4 | US-09-543-608A-39 |
| 11 | 89 | 52.7 | 21 | 5 | PCT-US95-02121-97 |
| 12 | 89 | 52.7 | 21 | 5 | PCT-US95-13841-20 |
| 13 | 86 | 50.9 | 423 | 2 | US-08-760-797A-1 |
| 14 | 86 | 50.9 | 424 | 2 | US-08-760-797A-3 |
| 15 | 86 | 50.9 | 424 | 3 | US-08-932-928B-1 |
| 16 | 86 | 50.9 | 424 | 3 | US-08-932-928B-3 |
| 17 | 82 | 48.5 | 412 | 1 | US-08-313-288B-18 |
| 18 | 79 | 46.7 | 17 | 4 | US-08-464-496-16 |
| 19 | 79 | 46.7 | 17 | 4 | US-08-197-484-96 |
| 20 | 79 | 46.7 | 17 | 5 | PCT-US95-02121-96 |
| 21 | 75 | 44.4 | 16 | 2 | US-08-817-933A-7 |
| 22 | 73.5 | 43.5 | 20 | 1 | US-08-465-167A-20 |
| 23 | 73.5 | 43.5 | 20 | 4 | US-08-627-820-20 |
| 24 | 73.5 | 43.5 | 20 | 5 | PCT-US92-07218-17 |
| 25 | 70 | 41.4 | 15 | 6 | US-09-100-414B-71 |
| 26 | 66 | 39.1 | 30 | 3 | US-09-303-323-71 |
| 27 | 66 | 39.1 | 30 | 3 | US-09-303-323-71 |

Sequence 71, Appl
Sequence 73, Appl
Sequence 73, Appl
Sequence 73, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 16, Appl
Sequence 12, Appl
Sequence 68, Appl
Sequence 68, Appl
Sequence 68, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 10, Appl

30 4 US-09-770-014-71
30 3 US-09-100-414B-73
30 3 US-09-303-323-73
30 4 US-09-770-014-73
16 1 US-08-465-167A-19
16 4 US-08-627-820-19
16 5 PCT-US92-07218-16
21 1 US-08-305-871A-12
30 3 US-09-100-414B-68
30 3 US-09-303-323-68
30 4 US-09-770-014-68
49 1 US-08-387-156-4
49 2 US-08-694-865-4
49 2 US-08-878-748-4
49 3 US-09-124-491-4
49 4 US-09-383-912-4
544 1 US-08-387-156-10
544 2 US-08-694-865-10

ALIGNMENTS

RESULT 1
US-08-446-692-27
; Sequence 27, Application US/08446692
; Patent No. 579551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-27

Query Match 79.6%; Score 134.5; DB 1; Length 33;
Best Local Similarity 90.3%; Pred. No. 3e-13;
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 2 EKIAKMEKASSVFNVNSGSLHWSYGLRP 32
DB 3 EKIAKMEKASSVFNVNSGSLHWSYGLRP 32

```

RESULT 2
US-08-488-351A-27      Application US/08488351A
; Sequence 27, Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-27

Query Match      79.6%; Score 134.5; DB 2; Length 33;
Best Local Similarity 90.3%; Pred. No. 3e-13;
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY      2 EKKIAKMEKASSVFNVNSGSLHWSYGLRP 32
DB      3 EKKIAKMEKASSVFNVNSGGE-HWSYGLRP 32

RESULT 3
US-08-186-266-6
; Sequence 6, Application US/08186266
; Patent No. 5662907
; GENERAL INFORMATION:
; APPLICANT: KUBO, Ralph T.
; APPLICANT: GREY, Howard M.
; APPLICANT: SETTE, Alessandro
; APPLICANT: CELIS, Esteban

```

```

; TITLE OF INVENTION: INDUCTION OF ANTI-TUMOR CYTOTOXIC
; TITLE OF INVENTION: T LYMPHOCYTES IN HUMANS USING
; TITLE OF INVENTION: SYNTHETIC PEPTIDE EPITOPES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,266
; FILING DATE: 25-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/159,339
; FILING DATE: 29-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-50-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1-21
; OTHER INFORMATION: /note= "Plasmodium falciparum CS
; OTHER INFORMATION: protein at positions 378-398."
US-08-186-266-6

Query Match      52.7%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAKMEKASSVFNVNS 20
DB      3 EKKIAKMEKASSVFNVNS 21

RESULT 4
US-08-446-692-48
; Sequence 48, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-446-692-48

Query Match 52.7%; Score 89; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0;

QY 2 EKXIAMEKASSVFVNVS 20
DB 3 EKXIAMEKASSVFVNVS 21

RESULT 5
US-08-488-351A-48
Sequence 48, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-48

Query Match 52.7%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0;

QY 2 EKXIAMEKASSVFVNVS 20
DB 3 EKXIAMEKASSVFVNVS 21

RESULT 6
US-09-100-409A-54
Sequence 54, Application US/09100409A
Patent No. 6090388
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
TITLE OF INVENTION: PREVENTION AND TREATMENT OF HIV INFECTION AND
TITLE OF INVENTION: IMMUNE DISORDERS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,409A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 1151-4154
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-409A-54

Query Match 52.7%; Score 89; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
 |||||
 Db 3 EKXIAKMEKASSVFNVNS 21

RESULT 7

US-08-464-496-17
 ; Sequence 17, Application US/08464496
 ; Patent No. 6322789
 ; GENERAL INFORMATION:
 ; APPLICANT: Epimmune, Inc.
 ; APPLICANT: Vitiello, Maria
 ; APPLICANT: Chesnut, Robert
 ; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
 ; TITLE OF INVENTION: EPITOPES
 ; FILE REFERENCE: 39963-20001.13
 ; CURRENT APPLICATION NUMBER: US/08/464,496
 ; PRIOR FILING DATE: 1995-06-05
 ; PRIOR APPLICATION NUMBER: 07/935,811
 ; PRIOR FILING DATE: 1992-08-26
 ; PRIOR APPLICATION NUMBER: 07/874,491
 ; PRIOR FILING DATE: 1992-04-27
 ; PRIOR APPLICATION NUMBER: 07/827,682
 ; PRIOR FILING DATE: 1992-01-29
 ; PRIOR APPLICATION NUMBER: 07/749,568
 ; PRIOR FILING DATE: 1991-08-26
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: Malaria circumsporozoite 378-398
 US-08-464-496-17

Query Match 52.7%; Score 89; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
 |||||
 Db 3 EKXIAKMEKASSVFNVNS 21

RESULT 8

US-08-788-822A-12
 ; Sequence 12, Application US/08788822A
 ; Patent No. 6413935
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexander, Jeffrey L.
 ; APPLICANT: Defrees, Shawn
 ; APPLICANT: Sette, Alessandro
 ; TITLE OF INVENTION: Induction of Immune Response Against
 ; TITLE OF INVENTION: Desired Determinants
 ; NUMBER OF SEQUENCES: 30
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/788,822A
 ; FILING DATE: 23-JAN-1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION INFORMATION:

APPLICATION NUMBER: US 60/010,510
 FILING DATE: 24-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Bastian, Kevin L.
 REGISTRATION NUMBER: 34,774
 REFERENCE/DOCKET NUMBER: 014137-009210US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-788-822A-12

Query Match 52.7%; Score 89; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
 |||||
 Db 3 EKXIAKMEKASSVFNVNS 21

RESULT 9

US-08-137-484-97
 ; Sequence 97, Application US/08197484
 ; Patent No. 6419931
 ; GENERAL INFORMATION:
 ; APPLICANT: VITIELLO, Maria A.
 ; APPLICANT: CHESTNUT, Robert W.
 ; APPLICANT: SETTE, Alessandro D.
 ; APPLICANT: CELIS, Esteban
 ; APPLICANT: GRAY, Howard
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
 ; TITLE OF INVENTION: CTL IMMUNITY
 ; NUMBER OF SEQUENCES: 153
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Khourie and Crew
 ; STREET: Steuart Street Tower, One Market Plaza
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94105-1493
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/197,484
 ; FILING DATE: 16-FEB-1994
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/935,811
 ; FILING DATE: 26-AUG-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/874,491
 ; FILING DATE: 27-APR-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/827,682
 ; FILING DATE: 29-JAN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/749,568
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parmelee, Steven W.
 ; REGISTRATION NUMBER: 31,990
 ; REFERENCE/DOCKET NUMBER: 14137-26-4
 ; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note="Malaria circumsporozoite
; OTHER INFORMATION: 378-398"
US-08-197-484-97

```

```

Query Match          52.7%; Score 89; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy  2 EKXIAMEKASSVFNVNS 20
Db  3 EKXIAMEKASSVFNVNS 21

```

RESULT 10

```

US-09-543-608A-39
; Sequence 39, Application US/09543608A
; Patent No. 6602510

```

GENERAL INFORMATION:

```

; APPLICANT: Fikes, John D.
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Celis, Esteban
; APPLICANT: Keogh, Elissa A.
; APPLICANT: Chesnut, Robert
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: HLA Class I A2 Tumor Associated Antigen
; TITLE OF INVENTION: Peptides and Vaccine Compositions
; FILE REFERENCE: 018623-015710US
; CURRENT APPLICATION NUMBER: US/09/543,608A
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmodium falciparum CS protein positions 378-398
US-09-543-608A-39

```

```

Query Match          52.7%; Score 89; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy  2 EKXIAMEKASSVFNVNS 20
Db  3 EKXIAMEKASSVFNVNS 21

```

RESULT 11

```

PCT-US95-02121-97
; Sequence 97, Application PC/TUS9502121
; GENERAL INFORMATION:

```

```

; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note="Malaria circumsporozoite
; OTHER INFORMATION: 378-398"
PCT-US95-02121-97

```

```

Query Match          52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy  2 EKXIAMEKASSVFNVNS 20
Db  3 EKXIAMEKASSVFNVNS 21

```

RESULT 12

```

PCT-US95-13841-20
; Sequence 20, Application PC/TUS9513841
; GENERAL INFORMATION:

```

```

; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IGE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lid, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-13841-20

Query Match 52.7%; Score 89; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNVS 20
Db 3 EKKIAKMEKASSVFNVNVS 21

RESULT 13
US-08-760-797A-1
; Sequence 1, Application US/08760797A
; Patent No. 5928302
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-760-797A-3

Query Match 50.9%; Score 86; DB 2; Length 424;
Best Local Similarity 67.7%; Pred. No. 9.4e-05;
Matches 21; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 2 EKKIAKMEKASSVFNVNVS 32
Db 171 EKKICKMEKCSVFNVNVS 195

RESULT 15
US-08-932-9298-1
; Sequence 1, Application US/089329298
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-760-797A-1

Query Match 50.9%; Score 86; DB 2; Length 423;
Best Local Similarity 85.7%; Pred. No. 9.3e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNVS 22
Db 174 EKKICKMEKCSVFNVNVS 194

RESULT 14
US-08-760-797A-3
; Sequence 3, Application US/08760797A
; Patent No. 5928302
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBSAG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-760-797A-3

Query Match 50.9%; Score 86; DB 2; Length 424;
Best Local Similarity 67.7%; Pred. No. 9.4e-05;
Matches 21; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 2 EKKIAKMEKASSVFNVNVS 32
Db 171 EKKICKMEKCSVFNVNVS 195

RESULT 15
US-08-932-9298-1
; Sequence 1, Application US/089329298
; Patent No. 6169171
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
```



```

;
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAg
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,929B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,797
; FILING DATE: 04-DEC-1996
; APPLICATION NUMBER: 08/442,612
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
; REGISTRATION NUMBER: 33,833
; REFERENCE/DOCKET NUMBER: B45015-1FWC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5096
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-932-929B-1

```

```

Query Match          50.9%; Score 86; DB 3; Length 424;
Best Local Similarity 85.7%; Pred. No. 9.4e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 EKKIAPMEKASVFNVNSGP 22
      |||||
Db      175 EKKICKMERKCSVFNVNSRP 195

```

```

Search completed: March 10, 2004, 09:28:54
Job time : 12.7763 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 25.9377 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKIAKKEKASSFNVNVSGLHWSVGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than, or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 168 | 99.4 | 33 | 9 | US-09-848-834A-12 |
| 2 | 157 | 92.9 | 51 | 9 | US-09-848-834A-20 |
| 3 | 95 | 56.2 | 20 | 9 | US-09-848-834A-3 |
| 4 | 95 | 56.2 | 36 | 9 | US-09-848-834A-16 |
| 5 | 92.5 | 54.7 | 31 | 15 | US-10-411-544-39 |
| 6 | 89 | 52.7 | 19 | 14 | US-10-239-313A-54 |
| 7 | 89 | 52.7 | 21 | 10 | US-09-932-165-1482 |
| 8 | 89 | 52.7 | 21 | 10 | US-09-935-384-710 |
| 9 | 89 | 52.7 | 21 | 10 | US-09-942-052-711 |
| 10 | 89 | 52.7 | 21 | 14 | US-10-001-469-1404 |
| 11 | 89 | 52.7 | 21 | 14 | US-10-128-711-97 |
| 12 | 89 | 52.7 | 21 | 14 | US-10-116-118-33 |
| 13 | 89 | 52.7 | 21 | 14 | US-10-062-109A-761 |
| 14 | 89 | 52.7 | 21 | 14 | US-10-005-480A-761 |
| 15 | 89 | 52.7 | 21 | 14 | US-10-277-292-652 |

| | | | | | | |
|----|------|------|-----|----|--------------------|---------------------|
| 16 | 89 | 52.7 | 21 | 15 | US-10-291-241-26 | Sequence 26, Appl |
| 17 | 89 | 52.7 | 21 | 15 | US-10-280-340-652 | Sequence 652, Appl |
| 18 | 89 | 52.7 | 21 | 15 | US-10-099-460-20 | Sequence 20, Appl |
| 19 | 89 | 52.7 | 21 | 15 | US-10-024-552-2583 | Sequence 2583, Appl |
| 20 | 89 | 52.7 | 21 | 15 | US-10-120-885A-25 | Sequence 25, Appl |
| 21 | 89 | 52.7 | 21 | 15 | US-10-107-532-2 | Sequence 2, Appl |
| 22 | 89 | 52.7 | 21 | 15 | US-10-121-016-52 | Sequence 52, Appl |
| 23 | 89 | 52.7 | 21 | 15 | US-10-114-669-2 | Sequence 44, Appl |
| 24 | 89 | 52.7 | 21 | 15 | US-10-120-835-44 | Sequence 4224, Appl |
| 25 | 89 | 52.7 | 21 | 15 | US-10-149-138-4224 | Sequence 39, Appl |
| 26 | 89 | 52.7 | 21 | 15 | US-10-114-432-39 | Sequence 19, Appl |
| 27 | 83.5 | 49.4 | 46 | 9 | US-09-848-834A-19 | Sequence 23, Appl |
| 28 | 82 | 48.5 | 21 | 10 | US-09-747-802-23 | Sequence 15, Appl |
| 29 | 82 | 48.5 | 21 | 10 | US-09-865-294-15 | Sequence 17, Appl |
| 30 | 82 | 48.5 | 21 | 15 | US-10-411-544-17 | Sequence 31, Appl |
| 31 | 82 | 48.5 | 396 | 10 | US-09-820-843A-31 | Sequence 11, Appl |
| 32 | 81.5 | 48.2 | 28 | 9 | US-09-848-834A-11 | Sequence 96, Appl |
| 33 | 79 | 46.7 | 17 | 14 | US-10-128-711-96 | Sequence 53, Appl |
| 34 | 79 | 46.7 | 17 | 14 | US-10-239-313A-53 | Sequence 3, Appl |
| 35 | 75 | 44.4 | 16 | 15 | US-10-372-111-3 | Sequence 36, Appl |
| 36 | 75 | 44.4 | 17 | 14 | US-10-239-313A-36 | Sequence 38, Appl |
| 37 | 75 | 44.4 | 17 | 14 | US-10-239-313A-38 | Sequence 39, Appl |
| 38 | 75 | 44.4 | 17 | 14 | US-10-239-313A-39 | Sequence 41, Appl |
| 39 | 75 | 44.4 | 17 | 14 | US-10-239-313A-41 | Sequence 60, Appl |
| 40 | 75 | 44.4 | 17 | 14 | US-10-239-313A-60 | Sequence 64, Appl |
| 41 | 75 | 44.4 | 17 | 14 | US-10-239-313A-64 | Sequence 67, Appl |
| 42 | 75 | 44.4 | 17 | 14 | US-10-239-313A-67 | Sequence 59, Appl |
| 43 | 74 | 43.8 | 17 | 14 | US-10-239-313A-59 | Sequence 65, Appl |
| 44 | 74 | 43.8 | 17 | 14 | US-10-239-313A-65 | Sequence 66, Appl |
| 45 | 74 | 43.8 | 17 | 14 | US-10-239-313A-66 | |

ALIGNMENTS

RESULT 1
US-09-848-834A-12
; Sequence 12, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of t
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino a
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated aspartic acid
; NAME/KEY: MOD_RES
; LOCATION: (33)..(33)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
; OTHER INFORMATION: (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (21)..(24)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (25)..(33)

RESULT 5
US-10-411-544-39
; Sequence 39, Application US/10411544
; Publication No. US20030232758A1
; GENERAL INFORMATION:
; APPLICANT: St. George-Hyslop, Peter
; APPLICANT: McLaurin, JoAnne
; TITLE OF INVENTION: Immunological Methods and Compositions for the Treatment of Alzheimer's Disease
; FILE REFERENCE: LI01547
; CURRENT APPLICATION NUMBER: US/10/411,544
; CURRENT FILING DATE: 2003-04-10
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: chimeric sequence
US-10-411-544-39

Query Match 54.7%; Score 92.5; DB 15; Length 31;
Best Local Similarity 79.3%; Pred. No. 1.7e-06;
Matches 23; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 2 EKXIAKMEKASSVFNVNNSGSLHWS-YG 29
Db 3 EKXIAKMEKASSVFNVNNSGGRHDSGYG 31

RESULT 6
US-10-239-313A-54
; Sequence 54, Application US/10239313A
; Publication No. US20030175285A1
; GENERAL INFORMATION:
; APPLICANT: KLINGUER - HAMOUR, Christine
; APPLICANT: CORVAIA, Nathalie
; APPLICANT: BECK, Alain
; APPLICANT: GOETSCH, Liliane
; TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM OF A GLUTAMATE
; FILE REFERENCE: 343 727 - US
; CURRENT APPLICATION NUMBER: US/10/239,313A
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: FR 00/03711
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT 01/70772
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 697
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Plasmodium malariae
US-10-239-313A-54

Query Match 52.7%; Score 89; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNNS 20
Db 1 EKXIAKMEKASSVFNVNNS 19

RESULT 7
US-09-932-165-1482
; Sequence 1482, Application US/09932165
; Publication No. US20030134784A1

RESULT 8
US-09-935-384-710
; Sequence 710, Application US/09935384
; Publication No. US2003016526A1
; GENERAL INFORMATION:
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: FARIS, MARY
; APPLICANT: GE, WANGMAO
; APPLICANT: JAKOBOVITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158P1H4
; FILE REFERENCE: 51158-20033.00
; CURRENT APPLICATION NUMBER: US/09/935,384
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 783
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 710
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-935-384-710

Query Match 52.7%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNNS 20
Db 3 EKXIAKMEKASSVFNVNNS 21

Query Match 52.7%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNNS 20
Db 3 EKXIAKMEKASSVFNVNNS 21

```
RESULT 9
US-09-942-052-711
; Sequence 711, Application US/09942052
; Publication No. US20030170626A1
; GENERAL INFORMATION:
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 711
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-942-052-711

Query Match      52.7%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAMEKASSVFNVNS 20
DB      3 EKKIAMEKASSVFNVNS 21

RESULT 10
US-10-001-469-1404
; Sequence 1404, Application US/10001469
; Publication No. US20030091562A1
; GENERAL INFORMATION:
; APPLICANT: JAKOBOVITS, AVA
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: SAFFRAN, DOUGLAS
; APPLICANT: HUBERT, RENE
; APPLICANT: FARIS, MARY
; APPLICANT: CHALLITA-EID, PIA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED
; TITLE OF INVENTION: 10192A11 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20024.20
; CURRENT APPLICATION NUMBER: US/10/001,469
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/157,902
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/291,118
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 09/680,728
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 2888
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1404
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-001-469-1404

Query Match      52.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAMEKASSVFNVNS 20
DB      3 EKKIAMEKASSVFNVNS 21
```

```
RESULT 11
US-10-128-711-97
; Sequence 97, Application US/10128711
; Publication No. US2003009634A1
; GENERAL INFORMATION:
; APPLICANT: VITIELLO, Maria A.
; APPLICANT: CHESTNUT, Robert W.
; APPLICANT: SETTE, Alessandro D.
; APPLICANT: CELIS, Esteban
; APPLICANT: GRAY, Howard
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/128,711
; FILING DATE: 22-Apr-2002
; CLASSIFICATION: Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION: /note= "Malaria circumsporozoite
; 378-398"
; SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-10-128-711-97

Query Match      52.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 EKKIAMEKASSVFNVNS 20
DB      3 EKKIAMEKASSVFNVNS 21
```

```
RESULT 12
US-10-116-118-33
; Sequence 761, Application US/10116118
; Publication No. US20030143672A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Ava
; TITLE OF INVENTION: Heteroclitic Analogs and Related Methods
; FILE REFERENCE: 2060.0090003
; CURRENT APPLICATION NUMBER: US/10/116,118
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/166,529
; PRIOR FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: US 60/239,008
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-116-118-33

Query Match      52.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKWEKASSVFNVNS 20
Db 3 EKKIAKWEKASSVFNVNS 21

RESULT 13
US-10-062-109A-761
; Sequence 761, Application US/10062109A
; Publication No. US20030165905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Ava
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20062.01
; CURRENT APPLICATION NUMBER: US/10/062,109A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 761
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-109A-761

Query Match      52.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKWEKASSVFNVNS 20
Db 3 EKKIAKWEKASSVFNVNS 21

RESULT 14
US-10-005-480A-761
; Sequence 761, Application US/10005480A
; Publication No. US20030191073A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Farris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Ava
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
; FILE REFERENCE: 51158-20062.00
; CURRENT APPLICATION NUMBER: US/10/005,480A
; CURRENT FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 761
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-005-480A-761

Query Match      52.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKWEKASSVFNVNS 20
Db 3 EKKIAKWEKASSVFNVNS 21

RESULT 15
US-10-277-292-652
; Sequence 652, Application US/10277292
; Publication No. US20030199470A1
; GENERAL INFORMATION:
; APPLICANT: FARRIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: AFAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOWITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PID7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/10/277,292
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/935,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 652
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-277-292-652

Query Match      52.7%; Score 89; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKWEKASSVFNVNS 20
Db 3 EKKIAKWEKASSVFNVNS 21
```

Wed Mar 10 10:34:15 2004

us-09-848-834a-12.open.rapb

Page 6

Search completed: March 10, 2004, 10:25:48
Job time : 25.9377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 49.6926 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-12

Perfect score: 169
Sequence: 1 DEKIAKMKASSVFNVNVGSLHWSYGLRPX 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 168 | 99.4 | 33 | 5 AAU11423 | AAU11423 Synthetic |
| 2 | 157 | 92.9 | 51 | 5 AAU11431 | AAU11431 Synthetic |
| 3 | 134.5 | 79.6 | 33 | 2 AAR62715 | AAR62715 LHRH-cont |
| 4 | 95 | 56.2 | 20 | 5 AAU11414 | AAU11414 P. falcip |
| 5 | 95 | 56.2 | 33 | 2 AAR83570 | AAR83570 IGE CH4 r |
| 6 | 95 | 56.2 | 36 | 5 AAU11427 | AAU11427 Synthetic |
| 7 | 89 | 52.7 | 19 | 4 AAR98951 | AAR98951 Vaccine r |
| 8 | 89 | 52.7 | 21 | 1 AAR91504 | AAR91504 Sequence |
| 9 | 89 | 52.7 | 21 | 2 AAR78920 | AAR78920 Malaria c |
| 10 | 89 | 52.7 | 21 | 2 AAR75955 | AAR75955 P. falcip |
| 11 | 89 | 52.7 | 21 | 2 AAR70912 | AAR70912 Malaria c |
| 12 | 89 | 52.7 | 21 | 2 AAR82586 | AAR82586 Plasmodiu |
| 13 | 89 | 52.7 | 21 | 2 AAR05612 | AAR05612 Circumspo |
| 14 | 89 | 52.7 | 21 | 2 AAR35440 | AAR35440 T-cell st |
| 15 | 89 | 52.7 | 21 | 2 AAY23252 | AAY23252 Peptide d |
| 16 | 89 | 52.7 | 21 | 3 AAV58777 | AAV58777 Unidentif |
| 17 | 89 | 52.7 | 21 | 3 AAV80071 | AAV80071 Pathogen |
| 18 | 89 | 52.7 | 21 | 3 AAV54553 | AAV54553 T helper |
| 19 | 89 | 52.7 | 21 | 4 AAR98457 | AAR98457 Plasmodiu |
| 20 | 89 | 52.7 | 21 | 4 AAR84447 | AAR84447 Sequence |
| 21 | 89 | 52.7 | 21 | 4 AAR84517 | AAR84517 Plasmodiu |
| 22 | 89 | 52.7 | 21 | 4 AAR99706 | AAR99706 Plasmodiu |
| 23 | 89 | 52.7 | 21 | 4 AAR88269 | AAR88269 Plasmodiu |
| 24 | 89 | 52.7 | 21 | 4 AAG62428 | AAG62428 Plasmodiu |
| 25 | 89 | 52.7 | 21 | 4 AAG89366 | AAG89366 Plasmodiu |

| | | | | | |
|----|----|------|----|------------|--------------------|
| 26 | 89 | 52.7 | 21 | 5 ABB94469 | ABB94469 Plasmodiu |
| 27 | 89 | 52.7 | 21 | 5 ABB78050 | ABB78050 Loosely M |
| 28 | 89 | 52.7 | 21 | 5 ABJ11372 | ABJ11372 P falcipa |
| 29 | 89 | 52.7 | 21 | 5 ABJ05781 | ABJ05781 P falcipa |
| 30 | 89 | 52.7 | 21 | 5 ABP51501 | ABP51501 Malaria c |
| 31 | 89 | 52.7 | 21 | 5 AAU95369 | AAU95369 Plasmodiu |
| 32 | 89 | 52.7 | 21 | 5 ABG34858 | ABG34858 P. falcip |
| 33 | 89 | 52.7 | 21 | 5 ABJ01952 | ABJ01952 158PID7 r |
| 34 | 89 | 52.7 | 21 | 5 AAU91545 | AAU91545 P. falcip |
| 35 | 89 | 52.7 | 21 | 5 ABJ16181 | ABJ16181 Zinc tran |
| 36 | 89 | 52.7 | 21 | 5 ABU72602 | ABU72602 Plasmodiu |
| 37 | 89 | 52.7 | 21 | 6 ABJ39281 | ABJ39281 Plasmodiu |
| 38 | 89 | 52.7 | 21 | 6 ABU98401 | ABU98401 Plasmodiu |
| 39 | 89 | 52.7 | 21 | 6 ABRO1863 | ABRO1863 Circumspo |
| 40 | 89 | 52.7 | 21 | 6 ABJ56921 | ABJ56921 Plasmodiu |
| 41 | 89 | 52.7 | 21 | 7 ADC71173 | ADC71173 P falcipa |
| 42 | 89 | 52.7 | 21 | 7 ADD84523 | ADD84523 Plasmodiu |
| 43 | 89 | 52.7 | 21 | 7 AD865914 | AD865914 Human 161 |
| 44 | 89 | 52.7 | 21 | 7 ADD96934 | ADD96934 P falcipa |
| 45 | 89 | 52.7 | 22 | 2 AAR82077 | AAR82077 Malaria C |

ALIGNMENTS

RESULT 1

AAU11423
ID AAU11423 standard; peptide; 33 AA.

XX AAU11423;

AC AC
DT 12-MAR-2002 (first entry)
XX Synthetic immunogen peptide 4.

DE Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
XX lutetinsing hormone releasing hormone; LHRH; contraceptive;

KW promiscuous helper T-cell peptide epitope; immunomic peptide epitope;
KW breast cancer; uterine cancer; gynecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Plasmodium falciparum.

OS Mammalia.

OS Synthetic.

OS Chimeric.

XX Key Location/Qualifiers

FT Peptide 1. .20

FT Peptide /note= "Malaria CSP protein (378-398 aa)"

FT Peptide 21. .24

FT Peptide /note= "Spacer peptide"

FT Peptide 25. .33

FT Modified-site 33 /note= "Gonadotrophin releasing hormone epitope"

FT /note= "Amidated glycine or glycine amide"

WO200185763-A2.

15-NOV-2001.

04-MAY-2001; 2001WO-US014363.

05-MAY-2000; 2000US-0202328P.

(APHT-) APHTON CORP.

Grimes S, Michaeli D, Stevens VC;

WPI; 2002-049440/06.

Novel synthetic immunogen for inducing immune response against
gonadotropin releasing hormone, comprises fusion peptide having
promiscuous helper T-cell peptide epitope and immunomic peptide epitope

| | |
|----------|---|
| XX | 05-MAY-2000; 2000US-0202328P. |
| PR | (APHT-) APHTON CORP. |
| XX | |
| PA | Grimes S, Michaeli D, Stevens VC; |
| XX | WPI; 2002-0494440/06. |
| PI | |
| XX | Novel synthetic immunogen for inducing immune response against |
| DR | gonadotropin releasing hormone, comprises fusion peptide having |
| XX | promiscuous helper T-cell peptide epitope and immunomimic peptide epitope |
| PT | or its analog. |
| XX | |
| XX | Claim 11; Page 12-13; 43pp; English. |
| PS | |
| XX | The invention relates to a synthetic immunogen for inducing specific |
| CC | antibodies against gonadotropin releasing hormone (GnRH) also known as |
| CC | lutinising hormone releasing hormone, LHRH) comprising a fusion peptide |
| CC | which comprises a promiscuous helper T-cell peptide epitope and |
| CC | immunomimic peptide epitope or its analogue. The synthetic immunogen is |
| CC | useful inducing an immune response against GnRH in an animal subject, and |
| CC | as such is useful as a contraceptive and in the treatment of diseases |
| CC | such as cancer (of the breast, uterus and other gynaecological cancer), |
| CC | endometriosis, uterine fibroids, benign prostatic hyper trophy and |
| CC | prostate cancer. The immunogen is effective in eliciting high and |
| CC | specific anti-GnRH antibody titres. The present sequence is a synthetic |
| CC | immunogen of the invention |
| XX | |
| SQ | Sequence 51 AA; |
| | |
| | Query Match 92.9%; Score 157; DB 5; Length 51; |
| | Best Local Similarity 94.1%; Pred. No. 2.3e-16; |
| | Matches 32; Conservative 0; Mismatches 0; Indels 2; Gaps 1; |
| | |
| Qy | 1 DEKKIAKWEKASSVFNVN--SGPSLHWSYGLRP 32 |
| | |
| Db | 17 DEKKIAKWEKASSVFNVNSSSGPSLHWSYGLRP 50 |
| | |
| | |
| RESULT 3 | |
| AAR62715 | |
| ID | AAR62715 standard; peptide; 33 AA. |
| AC | |
| XX | AAR62715; |
| AC | |
| XX | |
| DT | 25-MAR-2003 (revised) |
| DT | 10-SEP-1995 (first entry) |
| XX | |
| DE | LHRH-containing immunogenic peptide. |
| XX | |
| KW | Helper T cell epitope; universal immune stimulator; invasive; haptens; |
| KW | vaccine; LHRH; luteinising hormone releasing hormone; prostate; |
| KW | androgen-dependent carcinoma; antitumour; infertility; |
| KW | Plasmodium falciparum circumsporozoite. |
| XX | |
| OS | Synthetic. |
| XX | |
| PH | Key |
| FT | Domain |
| FT | 1..21 |
| FT | /note= "Plasmodium falciparum circumsporozoite helper T |
| FT | cell epitope" |
| FT | 24..33 |
| FT | /note= "LHRH haptens" |
| XX | |
| PN | WO9425060-A1. |
| XX | |
| PD | 10-NOV-1994. |
| XX | |
| PF | 28-APR-1994; 94WO-US004832. |
| XX | |
| PR | 27-APR-1993; 93US-00057166. |
| PR | 14-APR-1994; 94US-00229275. |

XX (LADD//) LADD A E.
PA (WANG//) WANG C Y.
PA (ZAMB//) ZAMB T.
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
DR Immunogenic luteinizing hormone releasing hormone peptide(s) - that
PT suppress LHRH activity in males and females.
XX Claim 8; Page 86; 213pp; English.
XX Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The stimulator
CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC immune response to the coupled peptide in members of a heterogeneous
CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC sequence from the invasive protein of Yersinia. Spacer amino acid
CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
CC domains and between the immune stimulator and hapten components. When the
CC hapten is LHRH, then optionally the invasin domain can be omitted from
CC the immune stimulator component. The present sequence represents an LHRH-
CC containing, invasin-free immunogenic peptide as above which can be used
CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
CC endometriosis, benign uterine tumours, recurrent functional ovarian
CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
CC cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX Sequence 33 AA;
SQ Query Match 79.6%; Score 134.5; DB 2; Length 33;
Best Local Similarity 90.3%; Pred. No. 3.8e-13;
Matches 28; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
OY 2 EKXIAKMEKASSVFNVNWSGSLHWSYGLRP 32
DB 3 EKXIAKMEKASSVFNVNWSGGE-HWSYGLRP 32
RESULT 4
AAU11414
ID AAU11414 standard; peptide; 20 AA.
XX AAU11414;
AC AAU11414;
DT 12-MAR-2002 (first entry)
DE P. falciparum circumsporozoite protein, CSP, peptide.
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
KW luteinizing hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer;
KW circumsporozoite protein; CSP.
XX Plasmodium falciparum.
OS Plasmodium falciparum.
XX WO200185763-A2.
PN WO200185763-A2.
XX 15-NOV-2001.
PD 15-NOV-2001.
XX 04-MAY-2001; 2001WO-US014363.
PF 04-MAY-2001; 2001WO-US014363.
XX 05-MAY-2000; 2000US-0202328P.
PR 05-MAY-2000; 2000US-0202328P.
XX (APHT-) APHTON CORP.
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX Novel synthetic immunogen for inducing immune response against
PT gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
PT or its analog.
XX Disclosure; Page 28; 43pp; English.
XX The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known as
CC luteinizing hormone releasing hormone, LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a peptide
CC from circumsporozoite protein, CSP, a promiscuous helper T-cell peptide
CC epitope used in the immunogen of the invention
XX Sequence 20 AA;
SQ Query Match 56.2%; Score 95; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DEKKIAKMEKASSVFNVNWS 20
DB 1 DEKKIAKMEKASSVFNVNWS 20
RESULT 5
AAR83570
ID AAR83570 standard; peptide; 33 AA.
XX AAR83570;
AC AAR83570;
DT 13-JUN-1996 (first entry)
DE IGE CH4 region contg. peptide immunogen for treating allergies.
XX IGE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
KW vaccine; allergy; antibody; constant heavy chain.
XX Synthetic.
OS Synthetic.
XX WO9526365-A1.
PN WO9526365-A1.
XX 05-OCT-1995.
PD 05-OCT-1995.
XX 24-MAR-1995; 95WO-US0003741.
PF 24-MAR-1995; 95WO-US0003741.
XX 28-MAR-1994; 94US-00218461.
PR 28-MAR-1994; 94US-00218461.
XX 25-OCT-1994; 94US-00328912.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY;
XX Wang CY;
XX WPI; 1995-351297/45.
XX Synthetic peptide-based immunogen contg. IGE CH4 peptide and helper T
PT cell epitope - useful for eliciting antibody prodn. for allergy
PT treatment.
XX Claim 5; Page 72; 87pp; English.
XX

CC AAR82592-R82600 and AAR83560-R83581 are peptide immunogens that are
 CC useful in vaccines for treating allergic reactions. In the immunogens, an
 CC IgE CH4 peptide is attached C-terminally to a series of amino acids
 CC including a helper T cell epitope. The immunogen may also opt. contain a
 CC fatty acid or fatty acid derivative, an invasin domain or alpha-NH2. The
 CC immunogen produces high titres of antibodies to the effector site in
 CC human IgE heavy chain (the CH4 domain peptide) which inhibit mast cell
 CC activation and reduce allergen-induced IgE prodn. The immunogens may be
 CC used in either a radially branching multimeric form or a linearly
 CC arranged monomeric form
 XX
 SQ Sequence 33 AA;

Query Match 56.2%; Score 95; DB 2; Length 33;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;

Qy 2 EKXIAKMEKASSVFNVSNG 21
 Db 3 EKXIAKMEKASSVFNVSNG 22

RESULT 6
 AAU11427
 ID AAU11427 standard; peptide; 36 AA.

AC AAU11427;
 DT 12-MAR-2002 (first entry)
 DE Synthetic immunogen peptide 8.

CC Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
 CC luteinising hormone releasing hormone; LHRH; contraceptive;
 CC promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 CC breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 CC uterine fibroid; benign prostatic hypertrophy; prostate cancer.

OS Plasmodium falciparum.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.

Key Location/Qualifiers
 FT Peptide 1..10
 FT /note= "Gonadotrophin releasing hormone epitope"
 FT Misc-difference 1
 FT /label= OTHER
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
 FT Peptide 11..16
 FT /note= "Spacer peptide"
 FT Peptide 17..36
 FT /note= "Malaria CSP protein (378-398 aa)"

XX WO200185763-A2.
 XX 15-NOV-2001.
 XX 04-MAY-2001; 2001WO-US014363.
 XX 05-MAY-2000; 2000US-0202328P.
 XX (APHT-) APHTON CORP.

PI Grimes S, Michaeli D, Stevens VC;
 DR WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against
 PT gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.

PS Claim 11; Page 10; 43pp; English.

CC The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention

SQ Sequence 36 AA;

Query Match 56.2%; Score 95; DB 5; Length 36;
 Best Local Similarity 100.0%; Pred. No. 4.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEKKIAKMEKASSVFNVSNS 20
 Db 17 DEKKIAKMEKASSVFNVSNS 36

RESULT 7
 AAM98951
 ID AAM98951 standard; peptide; 19 AA.

AC AAM98951;

DT 07-DEC-2001 (first entry)

DE Vaccine related MHC ligand peptide SEQ ID NO:54.

CC Glutamic acid; glutamine; vaccine; major histocompatibility complex; MHC;
 CC immunomodulator; antiallergic; endocrine; neuroprotectant; virucidal;
 CC bactericidal; antiparasitic; fungicidal; cytostatic; medicine;
 CC pharmaceutical; immune disorder; immune deficiency; autoimmune;
 CC hypersensitivity; allergy; graft rejection; infection; hormonal disorder;
 CC central nervous system disease; cancer; melanoma; anti-melanoma vaccine;
 CC human immunodeficiency virus.

OS Plasmodium malariae.

PN WO200170772-A2.

PD 27-SEP-2001.

PF 22-MAR-2001; 2001WO-FR000872.

PR 23-MAR-2000; 2000FR-00003711.

PA (FABR) FABRE MEDICAMENT SA PIERRE.

XX Klinguer-Hamour C, Corvaia N, Beck A, Goetsch L;

XX WPI; 2001-611470/70.

PT Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid.

PS Claim 9; Page 39; 149pp; French.

CC The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue in
 CC the form of an addition salt with a strong, physiologically acceptable
 CC acid (II). Also described are: (a) a pharmaceutical composition
 CC containing at least one (I); (b) a vaccine containing at least one (I)
 CC where this is a major histocompatibility complex (MHC) ligand (Ia); (c) a
 CC method for in vitro diagnosis of diseases associated with the presence of

CC (Ia); (d) a kit for method (c) that includes a (Ia); and (e) a process
 CC for preparing (I). (I) has immunomodulator, endocrine, antiallergic,
 CC neuroprotectant, virucidal, bactericidal, antiparasitic, fungicidal, and
 CC cytostatic activities. (I) are useful, in human or veterinary medicine,
 CC in pharmaceutical compositions (for treating immune disorders, e.g.
 CC immune deficiency, autoimmune states, hypersensitivity, allergy, graft
 CC rejection, infection, hormonal disorders and central nervous system
 CC diseases), also, where (I) is a MHC ligand (Ia), in vaccines for
 CC treatment or prevention of: (i) viral, bacterial, parasitic or fungal
 CC infections; or (ii) of cancers. A particular application is in anti-
 CC melanoma vaccines. (I) are also useful for in vitro diagnosis of diseases
 CC associated with interactions between MHC and (I), e.g. melanoma and human
 CC immunodeficiency virus infection. AAM9898 to AAM9952 represent peptides
 CC which can be used in pharmaceutical compounds from the present invention
 XX
 XX
 SQ Sequence 19 AA;

Query Match 52.7%; Score 89; DB 4; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAMEKASSVFNVNS 20
 Db 1 EKXIAMEKASSVFNVNS 19

RESULT 8
 AAP91504
 ID AAP91504 standard; peptide; 21 AA.

AC AAP91504;

DT 13-MAR-1992 (first entry)

DE Sequence of modified Plasmodium circumsporozoite T-cell epitope CS.T3.

XW Circumsporozoite peptide; T-cell epitope; immunogenic composition;
 XW vaccine.

OS Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 1..2 /note= "May be H-Asp-Ile, H-Ile, or H-"

FT Misc-difference 19..21

FT " /note= "May be Val-Asn-Ser-OH, Val-Asn-OH, Val-OH or -OH"

XX EP343460-A.

XX 29-NOV-1989.

XX 12-MAY-1989; 89EP-00108619.

XX 24-MAY-1988; 89GB-00012214.

XX (HOFF) HOFFMANN-LA ROCHE AG.

XX Sinigaglia F;

XX WPI; 1989-349561/48.

XX Modified Plasmodium CS peptide - used as a universally recognised T-cell
 FT epitope in vaccines to elicit an immune response against pathogenic
 FT agents.

XX Claim 1; Page 16; 23pp; English.

XX Peptide CS.T3 (residues 1-21) corresponds to residues 378-398 of the CS
 CC protein from P. falciparum but contains 2 Ala residues in place of the
 CC native protein's Cys residues at positions 384 and 389. Also claimed is
 CC AAP91504 (or modified forms, see FT) associated with an antigenic
 CC structure representing a B-cell epitope, pref. a multiple antigenic

CC peptide, esp. multimers of the repeat sequences NAMP present in P.
 CC falciparum CS protein
 XX
 SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAMEKASSVFNVNS 20
 Db 3 EKXIAMEKASSVFNVNS 21

RESULT 9

AAR78920
 ID AAR78920 standard; peptide; 21 AA.

XX AAR78920;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 27-MAR-1996 (first entry)

XX Malaria circumsporozoite 378-398 helper T lymphocyte epitope.

XX 378-398; cytotoxic T; CTL; epitope; helper T; HTL; cell; lymphocyte;
 KW antigens; treatment; disease prevention; tumours; viruses; parasites;
 KW malaria circumsporozoite; hepatitis B.

XX Plasmodium falciparum.

XX WO952317-A1.

XX 24-AUG-1995.

XX 16-FEB-1995; 95WO-US002121.

XX 16-FEB-1994; 94US-00197484.

XX (CYTE-) CYTEL CORP.

XX Vitello MA, Chesnut RW, Sette AD, Celis E, Grey H;

XX WPI; 1995-302545/39.

XX Compn. inducing cytotoxic T lymphocyte response to pref. viral,
 FT bacterial, parasitic or tumour antigens - useful in the treatment and
 FT prevention of diseases associated with the antigen e.g. hepatitis B.

XX Disclosure; Page 29; 109pp; English.

XX A compn. which induces a cytotoxic T lymphocyte (CTL) response to an
 CC antigen (Ag) in a mammal comprises a CTL Ag response inducing peptide
 CC and a lipid conjugated helper T cell inducing peptide (i.e. AAR78918-
 CC R78922). The compn. induces a CTL response to bacterial, viral or tumour
 CC Ags, and is therefore useful in the treatment and prevention of diseases
 CC associated with the Ag, e.g. hepatitis B. (Updated on 25-MAR-2003 to
 CC correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)
 XX

SQ Sequence 21 AA;

Query Match 52.7%; Score 89; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAMEKASSVFNVNS 20
 Db 3 EKXIAMEKASSVFNVNS 21

RESULT 10
 AAR75955

ID AAR75955 standard; peptide; 21 AA.
 XX AC AAR75955;
 XX DT 06-MAR-1996 (first entry)
 XX DE P. falciparum CS protein epitope residues 378-398.
 XX KW MAGE-3; melanoma antigen; vaccine; immune response; immunogenic peptide;
 XX KW cytotoxic T lymphocyte response; CTL; melanoma; breast cancer; antibody.
 XX OS Plasmodium falciparum.
 XX PN WO9519783-A1.
 XX PD 27-JUL-1995.
 XX PF 25-JAN-1995; 95WO-US001000.
 XX PR 25-JAN-1994; 94US-00186266.
 XX PA (CYTE-) CYTEL CORP.
 XX FI Kubo RT, Grey HM, Sette A, Celis E;
 XX DR WPI; 1995-269270/35.
 XX PT Immunogenic peptide(s) that induce immune response to cancer cells - that
 XX PT express a MAGE-3 protein peptide epitope used in vaccines or adoptive
 XX PT immunotherapy to induce cytotoxic T lymphocytes.
 XX PS Disclosure; Page 14; 44pp; English.
 XX CC AAR75942 is derived from the sequence of the melanoma antigen (MAGE-3)
 CC protein and can be used to elicit a primary cytotoxic T lymphocyte
 CC response against cells expressing MAGE-3. Synthetic peptides AAR75945-53
 CC can be used therapeutically to elicit CTL responses to melanoma, breast,
 CC colon, prostate, or other cells which express proteins with this epitope.
 CC The peptides have specific HLA-A1 binding capacity. The peptides can be
 CC also used in vaccines, esp. combined with peptides such as AAR75955-56,
 CC which are T-helper epitopes
 XX SQ Sequence 21 AA;
 Query Match 52.7%; Score 89; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 EKTIKVEKASSVFNVNS 20
 DB 3 EKTIKVEKASSVFNVNS 21
 RESULT 11
 AAR70912
 ID AAR70912 standard; protein; 21 AA.
 XX AC AAR70912;
 XX DT 27-AUG-2003 (revised)
 XX DT 25-MAR-2003 (revised)
 XX DT 09-OCT-1995 (first entry)
 XX DE Malaria circumsporozoite 378-398 T helper peptide.
 XX KW Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours;
 XX KW malaria circumsporozoite 378-398; T helper peptide.
 XX OS Plasmodium falciparum.
 XX PN WO9504542-A1.
 XX PD 16-FEB-1995.
 Query Match 52.7%; Score 89; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 EKTIKVEKASSVFNVNS 20
 DB 3 EKTIKVEKASSVFNVNS 21
 RESULT 12
 AAR82586
 ID AAR82586 standard; peptide; 21 AA.
 XX AC AAR82586;
 XX DT 13-JUN-1996 (first entry)
 XX DE Plasmodium falciparum circumsporozoite helper T cell epitope, PF.
 XX KW IGE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 XX KW vaccine; allergy; antibody; constant heavy chain.
 XX OS Plasmodium falciparum.
 XX PN WO9526365-A1.
 XX PD 05-OCT-1995.
 XX PF 24-MAR-1995; 95WO-US003741.
 XX PR 28-MAR-1994; 94US-00218461.
 XX PR 25-OCT-1994; 94US-00328912.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY;
 XX DR WPI; 1995-351297/45.
 XX PT Synthetic peptide-based immunogen contg. IGE CH4 peptide and helper T
 XX PT cell epitope - useful for eliciting antibody prodn. for allergy
 XX PT treatment.
 XX PS Claim 3; Page 23; 87pp; English.
 XX CC AAR82571-91 are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IGE CH4 peptide is

XX 02-AUG-1994; 94WO-US008721.
 XX 06-AUG-1993; 93US-00103623.
 XX (CYTE-) CYTEL CORP.
 XX FI Fikes JD, Livingston BD, Sette AD, Sidney JC;
 XX DR WPI; 1995-090681/12.
 XX PT Human melanoma antigen, MAGE-1, peptide(s) - useful for stimulating
 XX PT immune response against melanoma.
 XX PS Disclosure; Page 13; 59pp; English.
 XX CC The T helper peptides described in AAR70910-R70914 are used in
 CC conjunction with the C-terminal MAGE-1 peptides described in AAR70915 to
 CC AAR70969. Compsns. containing the T helper and MAGE-1 peptides can be
 CC administered, as a vaccine to patients susceptible to MAGE associated
 CC tumours, e.g. melanomas. (Updated on 25-MAR-2003 to correct PN field.)
 XX CC (Updated on 27-AUG-2003 to correct OS field.)
 XX SQ Sequence 21 AA;
 Query Match 52.7%; Score 89; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 EKTIKVEKASSVFNVNS 20
 DB 3 EKTIKVEKASSVFNVNS 21
 RESULT 12
 AAR82586
 ID AAR82586 standard; peptide; 21 AA.
 XX AC AAR82586;
 XX DT 13-JUN-1996 (first entry)
 XX DE Plasmodium falciparum circumsporozoite helper T cell epitope, PF.
 XX KW IGE; CH4; immunoglobulin; epsilon; immunogen; helper T cell; epitope;
 XX KW vaccine; allergy; antibody; constant heavy chain.
 XX OS Plasmodium falciparum.
 XX PN WO9526365-A1.
 XX PD 05-OCT-1995.
 XX PF 24-MAR-1995; 95WO-US003741.
 XX PR 28-MAR-1994; 94US-00218461.
 XX PR 25-OCT-1994; 94US-00328912.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY;
 XX DR WPI; 1995-351297/45.
 XX PT Synthetic peptide-based immunogen contg. IGE CH4 peptide and helper T
 XX PT cell epitope - useful for eliciting antibody prodn. for allergy
 XX PT treatment.
 XX PS Claim 3; Page 23; 87pp; English.
 XX CC AAR82571-91 are helper T cell epitopes which can be used in the
 CC preparation of a peptide immunogen that is useful in vaccines for
 CC treating allergic reactions. In the immunogen an IGE CH4 peptide is

CC attached C-terminally to a series of amino acids including a helper T
 CC cell epitope. The immunogen may also opt. contain a fatty acid or fatty
 CC acid derivative, an invasion domain or alpha-NH2. The immunogen produces
 CC high titres of antibodies to the effector site in human IGE heavy chain
 CC (the CH4 domain peptide) which inhibit mast cell activation and reduce
 CC allergen-induced IGE prodn. The immunogens may be used in either a
 CC radially branching multimeric form or a linearly arranged monomeric form

XX Sequence 21 AA;

Query Match 52.7%; Score 89; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
 |||||
 DB 3 EKXIAKMEKASSVFNVNS 21

RESULT 13

AAW05612
 ID AAW05612 standard; peptide; 21 AA.

XX AC AAW05612;

DT 10-DEC-1996 (first entry)

DE Circumsporozoite helper T cell epitope.

XX Immunoglobulin; IGE; membrane protein; human; epsilon chain; hepatitis B;
 KW membrane anchoring domain; helper T cell; surface antigen; core antigen;
 KW pertussis toxin; tetanus toxin; measles virus F protein; immunotherapy;
 KW Chlamydia trachomatis major outer membrane protein; immunogen; vaccine;
 KW diphtheria toxin; Plasmodium falciparum; circumsporozoite; E. coli Trt;
 KW schistosoma mansoni; triose phosphate isomerase; allergenic reaction;
 KW allergic rhinitis; food allergy; anaphylaxis; virally-induced asthma;
 KW antihistamine; decongestant; beta-2 agonist; immunosuppression;
 KW corticosteroid.

XX Plasmodium falciparum.

XX WO9612740-A1.

PN 02-MAY-1996.

PF 25-OCT-1995; 95WO-US013841.

PR 25-OCT-1994; 94US-00328519.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 1996-230555/23.

XX Peptide immunogen useful in treatment of allergy - comprises membrane-
 PT bound IGE epsilon-chain peptide synthesised linearly in tandem with T
 PT helper epitope peptide.

PS Claim 2; Page 19; 53pp; English.

XX AAW05957-W05616 represent helper T cell epitopes used in the peptide
 CC immunogens of the invention. This sequence represents the Plasmodium
 CC falciparum circumsporozoite helper T cell antigen. The peptides of the
 CC invention contain one of these sequences, and a membrane-bound
 CC immunoglobulin E (IGE) fragment (see AAW05595 and AAW05596). The peptide
 CC immunogens of the invention can be used in vaccines for the
 CC immunotherapeutic treatment of allergenic reactions, including allergic
 CC rhinitis, food allergies, anaphylaxis, or virally-induced asthma. The
 CC immunogens overcome the short effective period of antihistamines,
 CC decongestants, and beta-2 agonists, while preventing the broad
 CC immunosuppression of corticosteroids. The peptides do not have the
 CC potential side effects of restlessness or sedation (associated with

CC antihistamines), associated increased morbidity in asthmatics (as seen
 CC with beta-2 agonists) and adverse hormonal activities (observed in
 CC corticosteroid users)

XX Sequence 21 AA;

Query Match 52.7%; Score 89; DB 2; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASSVFNVNS 20
 |||||
 DB 3 EKXIAKMEKASSVFNVNS 21

RESULT 14

AAW35440

ID AAW35440 standard; peptide; 21 AA.

XX AC AAW35440;

XX 25-MAR-2003 (revised)

DT 22-APR-1998 (first entry)

DE T-cell stimulatory peptide from Plasmodium falciparum.

XX T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
 KW scaffold; inhibition; metastasis; wound healing; solid phase.

XX Plasmodium falciparum.

XX WO9738011-A1.

PD 16-OCT-1997.

PF 03-APR-1997; 97WO-DK000146.

PR 03-APR-1996; 96DK-00000398.

XX (PEPR-) PEPRESEARCH AS.

XX Heegaard PMH, Jakobsen PH;

XX WPI; 1997-512645/47.

XX Non-dendritic peptide carrier linked to a solid phase - useful as a
 PT diagnostic agent and as a scaffold for production of chemical
 PT derivatives.

PS Claim 30; Page 199; 262pp; English.

XX A non-dendritic peptide carrier (A) has been developed which is coupled
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.
 CC Where (A) comprises 10-50 amino acids capable of forming a secondary
 CC structure in a benign buffer after liberation from the solid phase, and
 CC further the (A)-solid phase complex comprises an immunogenic substance
 CC and/or an immune mediator coupled on (A). The present sequence represents
 CC a specifically claimed T-cell stimulatory peptide from the present
 CC invention. An (A)-solid phase complex can be used as a scaffold for the
 CC production of chemical derivatives, characterised by covalently attaching
 CC molecules at attachment points. Alternatively (A) is used as a scaffold-
 CC peptide for the incorporation into an immunostimulating complex (Iscom)
 CC resulting in an (A)-Iscom complex which is used for the chemical coupling of
 CC antigenic substances in an aqueous solution by conjugation. (A)
 CC derivatised with one or more peptides having fibronectin-, laminin- or
 CC vitronectin-like binding activities can be used for the promotion of cell
 CC attachment to plastic surfaces, in particular to inhibit tumour growth
 CC and metastasis, and for promotion of wound healing. Also a derivatised
 CC (A) can be used for the selection of specifically-binding aptamers or as
 CC a diagnostic agent. Such diagnostic-(A) molecules could be used to detect
 CC molecules derived from or indicative of pregnancy or of a disease, such
 CC as an infectious, autoimmune or cancerous disease. (Updated on 25-MAR-
 CC 2003 to correct PF field.)


```

XX SQ Sequence 21 AA;
Query Match 52.7%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVVNS 20
Db 3 EKKIAKMEKASSVFNVVNS 21

RESULT 15
AAV23252
ID AAY23252 standard; peptide; 21 AA.
XX AC AAY23252;
XX DT 31-AUG-1999 (first entry)
XX DE Peptide derived from Plasmodium falciparum CS protein.
XX KW Venezuelan equine encephalitis virus; VEE virus; neoplastic disease;
XX KW tumour-associated antigen; cytokine; immunity; cancer; tumour;
XX KW CS protein.
XX OS Plasmodium falciparum.
XX PN WO9930734-A1.
XX PD 24-JUN-1999.
XX PF 14-DEC-1998; 98WO-US025725.
XX PR 18-DEC-1997; 97US-0068080P.
XX PA (SEAR ) SEARLE & CO G D.
XX PI Hippenmeyer PJ;
XX DR WPI; 1999-395093/33.
XX PT Using new Venezuelan equine encephalitis virus vectors.
XX PS Claim 5; Page 24; 40pp; English.
XX CC The specification describes Venezuelan equine encephalitis (VEE) virus
CC vectors which can be used to express tumour-associated antigens and
CC cytokines, and thus induce immunity to cancer. The VEE virus vectors of
CC the invention can be used prevent, treat, and protect against primary and
CC metastatic neoplastic diseases, especially tumours such as lung cancer,
CC breast cancer, ovarian cancer, prostate cancer, pancreatic cancer,
CC gastric cancer, colon cancer, renal cancer, bladder cancer, melanoma,
CC hepatoma, sarcoma and lymphoma. The present sequence is derived from the
CC Plasmodium falciparum CS protein, and can be fused with the target
CC peptide of the invention to provide a greater stimulation of the immune
CC system
XX SQ Sequence 21 AA;
Query Match 52.7%; Score 89; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVVNS 20
Db 3 EKKIAKMEKASSVFNVVNS 21

Search completed: March 10, 2004, 09:12:11
Job time : 50.6926 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 32.6148 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKIAKMEKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|-------------------|
| 1 | 82 | 48.5 | 69 | 5 Q8MZK7 | Q8MZK7 plasmodium |
| 2 | 82 | 48.5 | 115 | 5 Q25837 | Q25837 plasmodium |
| 3 | 82 | 48.5 | 115 | 5 Q25839 | Q25839 plasmodium |
| 4 | 82 | 48.5 | 115 | 5 Q9U934 | Q9U934 plasmodium |
| 5 | 82 | 48.5 | 115 | 5 Q25835 | Q25835 plasmodium |
| 6 | 82 | 48.5 | 115 | 5 Q25836 | Q25836 plasmodium |
| 7 | 82 | 48.5 | 115 | 5 Q9U9V7 | Q9U9V7 plasmodium |
| 8 | 82 | 48.5 | 117 | 5 Q25795 | Q25795 plasmodium |
| 9 | 82 | 48.5 | 117 | 5 Q25796 | Q25796 plasmodium |
| 10 | 82 | 48.5 | 117 | 5 Q25797 | Q25797 plasmodium |
| 11 | 82 | 48.5 | 117 | 5 Q25794 | Q25794 plasmodium |
| 12 | 82 | 48.5 | 356 | 5 Q819G5 | Q819G5 plasmodium |
| 13 | 82 | 48.5 | 360 | 5 Q819H7 | Q819H7 plasmodium |
| 14 | 82 | 48.5 | 360 | 5 Q819H6 | Q819H6 plasmodium |
| 15 | 82 | 48.5 | 363 | 5 Q819H3 | Q819H3 plasmodium |
| 16 | 82 | 48.5 | 364 | 5 Q819G4 | Q819G4 plasmodium |

| | | | | | |
|----|----|------|-----|----------|-------------------|
| 17 | 82 | 48.5 | 368 | 5 Q819G3 | Q819G3 plasmodium |
| 18 | 82 | 48.5 | 380 | 5 Q819G8 | Q819G8 plasmodium |
| 19 | 82 | 48.5 | 380 | 5 Q819G1 | Q819G1 plasmodium |
| 20 | 82 | 48.5 | 385 | 5 Q819I9 | Q819I9 plasmodium |
| 21 | 82 | 48.5 | 385 | 5 Q819I8 | Q819I8 plasmodium |
| 22 | 82 | 48.5 | 388 | 5 Q819H4 | Q819H4 plasmodium |
| 23 | 82 | 48.5 | 389 | 5 Q819J2 | Q819J2 plasmodium |
| 24 | 82 | 48.5 | 389 | 5 Q81056 | Q81056 plasmodium |
| 25 | 82 | 48.5 | 392 | 5 Q819H9 | Q819H9 plasmodium |
| 26 | 82 | 48.5 | 392 | 5 Q819H1 | Q819H1 plasmodium |
| 27 | 82 | 48.5 | 392 | 5 Q819H0 | Q819H0 plasmodium |
| 28 | 82 | 48.5 | 392 | 5 Q819G2 | Q819G2 plasmodium |
| 29 | 82 | 48.5 | 393 | 5 Q819J1 | Q819J1 plasmodium |
| 30 | 82 | 48.5 | 393 | 5 Q819J0 | Q819J0 plasmodium |
| 31 | 82 | 48.5 | 393 | 5 Q819I1 | Q819I1 plasmodium |
| 32 | 82 | 48.5 | 396 | 5 Q819I0 | Q819I0 plasmodium |
| 33 | 82 | 48.5 | 396 | 5 Q819H9 | Q819H9 plasmodium |
| 34 | 82 | 48.5 | 396 | 5 Q819I7 | Q819I7 plasmodium |
| 35 | 82 | 48.5 | 397 | 5 Q819I6 | Q819I6 plasmodium |
| 36 | 82 | 48.5 | 404 | 5 Q819I5 | Q819I5 plasmodium |
| 37 | 82 | 48.5 | 408 | 5 Q25729 | Q25729 plasmodium |
| 38 | 82 | 48.5 | 408 | 5 Q819G9 | Q819G9 plasmodium |
| 39 | 82 | 48.5 | 408 | 5 Q810I0 | Q810I0 plasmodium |
| 40 | 82 | 48.5 | 412 | 5 Q819I4 | Q819I4 plasmodium |
| 41 | 82 | 48.5 | 412 | 5 Q819H5 | Q819H5 plasmodium |
| 42 | 82 | 48.5 | 412 | 5 Q819H2 | Q819H2 plasmodium |
| 43 | 82 | 48.5 | 412 | 5 Q819G7 | Q819G7 plasmodium |
| 44 | 82 | 48.5 | 412 | 5 Q819G5 | Q819G5 plasmodium |
| 45 | 82 | 48.5 | 412 | 5 Q810R5 | Q810R5 plasmodium |

ALIGNMENTS

RESULT 1
Q8MZK7 PRELIMINARY; PRT; 69 AA.
AC Q8MZK7; 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein II (Fragment).
GN CSPII.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC1/HN;
RA Chen H.H., Yu X.B., Wu Z.D., Xu J.;
RT "Gene cloning of circumsporozoite protein (CSP) II gene from
RT Plasmodium falciparum (FC1/HN).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY093672; AAM19072.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IBA.
DR InterPro; IPR0003067; Crcmsprzoite.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF00090; tsd.1; 1.
DR PRINTS; PRO3303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI; 1.
DR PROSITE; PSS0092; TSPI; 1.
FT NON TER 1 1
SQ SEQUENCE 69 AA; 7670 MW; F3E1AA54671FE6FF CRC64;

Query Match 48.5%; Score 82; DB 5; Length 69;
Best Local Similarity 78.3%; Pred No. 0.00028;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 EKIAKMEKASSVFNVNSGSL 24
DB 37 EKICKMEKCKSSVFNVNSIGL 59

```
RESULT 2
Q25837      PRELIMINARY;      PRT; 115 AA.
AC Q25837;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835a;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RI falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83160; AAA29573.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI.1.
DR PROSITE; PS50092; TSPI.1.
FT NON TER 1
SQ SEQUENCE 115 AA; 13018 MW; COA23F5805688237 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSVFVNVSIGL 105

RESULT 3
Q25839      PRELIMINARY;      PRT; 115 AA.
AC Q25839;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835c;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RI falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83162; AAA29575.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSPI.1.
DR PROSITE; PS50092; TSPI.1.
FT NON TER 1
SQ SEQUENCE 115 AA; 12974 MW; 6B5538287260DA90 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSVFVNVSIGL 105

RESULT 4
Q25835      PRELIMINARY;      PRT; 115 AA.
AC Q25835;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=828;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RI falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83157; AAA29570.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
FT NON TER 1
SQ SEQUENCE 115 AA; 12925 MW; A24CB33ABEE3F232 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSVFVNVSIGL 105

RESULT 5
Q25835      PRELIMINARY;      PRT; 115 AA.
AC Q25835;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=828;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RI falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83157; AAA29570.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
FT NON TER 1
SQ SEQUENCE 115 AA; 12925 MW; A24CB33ABEE3F232 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSVFVNVSIGL 105
```

```
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; 9BES2CPFF12BD48F CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSSVFVNVSIGL 105

RESULT 6
Q25836 PRELIMINARY; PRT; 115 AA.
AC Q25836;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=834b;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwattiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates."
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83159; AAA29572.1; -.
DR InterPro; IPR003793; F:defense/immunity protein activity; IEA.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12926 MW; 697DF630E50B3A30 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSSVFVNVSIGL 105

RESULT 7
Q25796 PRELIMINARY; PRT; 115 AA.
AC Q25796;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=815, 947, and 808;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwattiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates."
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83153; AAA29566.1; -.
DR EMBL; M83171; AAA29549.1; -.
DR EMBL; M83151; AAA29564.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 115 AA; 12925 MW; 697DF63EBBBD90 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 115;
Best Local Similarity 78.3%; Pred. No. 0.00049;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 83 EKKICKMEKCSSVFVNVSIGL 105

RESULT 8
Q25795 PRELIMINARY; PRT; 117 AA.
AC Q25795;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Doonan D.L., Saul A., Good M.F.;
RT "Geographically restricted heterogeneity of the Plasmodium falciparum
circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M77203; AAA29517.2; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
FT NON_TER 1
SQ SEQUENCE 117 AA; 13110 MW; D96BE20944A7C726 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 117;
Best Local Similarity 78.3%; Pred. No. 0.0005;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFVNVSIGL 24
Db 85 EKKICKMEKCSSVFVNVSIGL 107

RESULT 9
Q25796 PRELIMINARY; PRT; 117 AA.
AC Q25796;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
```

```

RP SEQUENCE FROM N.A.
RA Doonan D.L., Saul A., Good M.F.;
RT "Geographically restricted heterogeneity of the Plasmodium falciparum
RT Circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBSJ databases.
DR EMBL; M77204; AAA29518.2; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PSS0092; TSP1; 1.
FT NON TER 1
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 117 AA; 13052 MW; D96BE209420BADE6 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 117;
Best Local Similarity 78.3%; Pred. No. 0.0005; 5; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
Db 85 EKKICKMEKCSVFNVNNGSIGL 107

RESULT 10
ID Q25797 PRELIMINARY; PRT; 117 AA.
AC Q25797;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CX NCBI_TaxID=5833;
RN Doonan D.L., Saul A., Good M.F.;
RA "Geographically restricted heterogeneity of the Plasmodium falciparum
RT Circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBSJ databases.
DR EMBL; M77205; AAA29519.2; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PSS0092; TSP1; 1.
FT NON TER 1
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 117 AA; 13043 MW; ODA711D86C0B03C1 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 117;
Best Local Similarity 78.3%; Pred. No. 0.0005; 5; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
Db 85 EKKICKMEKCSVFNVNNGSIGL 107

RESULT 11
ID Q25794 PRELIMINARY; PRT; 117 AA.
AC Q25794;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Circumsporozoite protein (Fragment).
OS Plasmodium falciparum.

```

```

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN Doonan D.L., Saul A., Good M.F.;
RA "Geographically restricted heterogeneity of the Plasmodium falciparum
RT Circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBSJ databases.
DR EMBL; M77202; AAA29516.2; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PSS0092; TSP1; 1.
FT NON TER 1
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 117 AA; 13080 MW; 8F9B92CC10ECC212 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 117;
Best Local Similarity 78.3%; Pred. No. 0.0005; 5; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
Db 85 EKKICKMEKCSVFNVNNGSIGL 107

RESULT 12
ID Q819G5 PRELIMINARY; PRT; 356 AA.
AC Q819G5;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum (isolate fid3 / India).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=70152;
RN Doonan D.L., Saul A., Good M.F.;
RA "Geographically restricted heterogeneity of the Plasmodium falciparum
RT Circumsporozoite protein: relevance for vaccine development.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBSJ databases.
DR EMBL; M77203; AAA29517.2; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PSS0092; TSP1; 1.
SQ SEQUENCE 356 AA; 38840 MW; B1AB87D66609A100 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 356;
Best Local Similarity 78.3%; Pred. No. 0.0017; 5; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNGPSL 24
Db 324 EKKICKMEKCSVFNVNNGSIGL 346

RESULT 13
ID Q819H7 PRELIMINARY; PRT; 360 AA.
AC Q819H7;

```

DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Circumsporozoite protein.
 GN CSP.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Asemo Bay;
 RX MEDLINE=22356746; PubMed=12467976;
 RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
 RA Magris M., Biswas S., Kariuki S., Lal A.A.;
 RT "A study of genetic diversity in the gene encoding the
 RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
 RT transmission areas-XVI. Asemo Bay Cohort Project.";
 RL Mol. Biochem. Parasitol. 125:83-90(2002).
 DR EMBL; AF540462; AAN87596.1; -
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Crcmsprzoite.
 DR Pfam; PF00090; tsp.1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50092; TSP1; 1.
 SQ SEQUENCE 360 AA; 39093 MW; 6504CC012649236C CRC64;

Query Match 48.5%; Score 82; DB 5; Length 360;
 Best Local Similarity 78.3%; Pred. No. 0.0018; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKXIARMEKASSVFNVNSGPSL 24
 |||||
 Db 328 EKXICKMEKCSSVFNVNSIGL 350

RESULT 14

Q819H6 PRELIMINARY; PRT; 360 AA.
 AC Q819H6;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Circumsporozoite protein.
 GN CSP.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Asemo Bay;
 RX MEDLINE=22356746; PubMed=12467976;
 RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
 RA Magris M., Biswas S., Kariuki S., Lal A.A.;
 RT "A study of genetic diversity in the gene encoding the
 RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
 RT transmission areas-XVI. Asemo Bay Cohort Project.";
 RL Mol. Biochem. Parasitol. 125:83-90(2002).
 DR EMBL; AF540463; AAN87597.1; -
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Crcmsprzoite.
 DR Pfam; PF00090; tsp.1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50092; TSP1; 1.
 SQ SEQUENCE 360 AA; 39062 MW; 65058844270D666C CRC64;

Query Match 48.5%; Score 82; DB 5; Length 360;
 Best Local Similarity 78.3%; Pred. No. 0.0018; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKXIARMEKASSVFNVNSGPSL 24
 |||||
 Db 328 EKXICKMEKCSSVFNVNSIGL 350

RESULT 15

Q819H3 PRELIMINARY; PRT; 363 AA.
 AC Q819H3;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Circumsporozoite protein.
 GN CSP.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Boliviar;
 RX MEDLINE=22356746; PubMed=12467976;
 RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
 RA Magris M., Biswas S., Kariuki S., Lal A.A.;
 RT "A study of genetic diversity in the gene encoding the
 RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
 RT transmission areas-XVI. Asemo Bay Cohort Project.";
 RL Mol. Biochem. Parasitol. 125:83-90(2002).
 DR EMBL; AF540466; AAN87600.1; -
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR InterPro; IPR003067; Crcmsprzoite.
 DR Pfam; PF00090; tsp.1; 1.
 DR PRINTS; PR01303; CRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50092; TSP1; 1.
 SQ SEQUENCE 363 AA; 39458 MW; E88E17AC5A11BD55 CRC64;

Query Match 48.5%; Score 82; DB 5; Length 363;
 Best Local Similarity 78.3%; Pred. No. 0.0018; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKXIARMEKASSVFNVNSGPSL 24
 |||||
 Db 331 EKXICKMEKCSSVFNVNSIGL 353

Search completed: March 10, 2004, 09:25:35
 Job time : 33.6148 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 6.03502 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKXIAQKEXKASSVFNVNSGSLHWSYGLRPX 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|--------------|---------------------|
| 1 | 82 | 48.5 | 388 | 1 CSP_PLARE | P26694 plasmodium |
| 2 | 82 | 48.5 | 397 | 1 CSP_PLAFO | P19597 Plasmodium |
| 3 | 82 | 48.5 | 412 | 1 CSP_PLAFA | P02893 Plasmodium |
| 4 | 82 | 48.5 | 424 | 1 CSP_PLAFT | P13814 plasmodium |
| 5 | 82 | 48.5 | 442 | 1 CSP_PLAFW | P08307 plasmodium |
| 6 | 56.5 | 33.4 | 92 | 1 GON1_RAT | P07490 rattus norv |
| 7 | 55 | 32.5 | 339 | 1 CSP_PLABE | P06915 plasmodium |
| 8 | 55 | 32.5 | 347 | 1 CSP_PLABA | P23093 Plasmodium |
| 9 | 55 | 32.5 | 367 | 1 CSP_PLAYO | P05914 plasmodium |
| 10 | 54 | 32.0 | 67 | 1 GON1_MACMU | P55247 macaca mula |
| 11 | 54 | 32.0 | 90 | 1 GON1_MOUSE | P13562 mus musculus |
| 12 | 54 | 32.0 | 91 | 1 GON1_PIG | P49921 sus scrofa |
| 13 | 54 | 32.0 | 92 | 1 GON1_HUMAN | P01148 homo sapien |
| 14 | 54 | 32.0 | 95 | 1 GON1_PAGMA | P70074 pagrus majo |
| 15 | 52 | 30.8 | 61 | 1 GON1_SHEEP | Q28588 ovis aries |
| 16 | 52 | 30.8 | 63 | 1 GON1_VESAU | O09163 mesocricetu |
| 17 | 52 | 30.8 | 89 | 1 GON1_XENLA | P45656 xenopus lae |
| 18 | 52 | 30.8 | 90 | 1 GON1_RANCA | Q90y63 rana catesb |
| 19 | 52 | 30.8 | 92 | 1 GON1_TUPGS | Q95335 tupiaa glis |
| 20 | 52 | 30.8 | 994 | 1 DPO2_KULLA | P05468 kluyveromyc |
| 21 | 51 | 30.2 | 91 | 1 GON1_ORYLA | Q94gdc o prognado |
| 22 | 51 | 30.2 | 92 | 1 GON1_CAVPO | O54713 cavia porce |
| 23 | 51 | 30.2 | 393 | 1 CSP_PLABR | P14593 plasmodium |
| 24 | 51 | 30.2 | 429 | 1 CSP_PLAWA | P13815 plasmodium |
| 25 | 51 | 30.2 | 473 | 1 SYE_WIGBR | Q86375 wigglewort |
| 26 | 50 | 29.6 | 95 | 1 GON1_SPAAU | P51919 sparus aura |
| 27 | 50 | 29.6 | 2339 | 1 RPCI_PLAFA | P27625 plasmodium |
| 28 | 49.5 | 29.3 | 630 | 1 ESA8_TRYBB | P23799 trypanosoma |
| 29 | 49.5 | 29.3 | 3176 | 1 CA36_HUMAN | P12111 homo sapien |
| 30 | 49 | 29.0 | 95 | 1 GON1_MORSA | O73812 morone saxa |
| 31 | 49 | 29.0 | 532 | 1 INV4_YEAST | P10596 saccharomyc |
| 32 | 48.5 | 28.0 | 1146 | 1 ITAL_DROME | Q24247 drosophila |
| 33 | 48.5 | 28.7 | 630 | 1 ESA8_TRYEQ | P26337 trypanosoma |

34 48 28.4 10 1 GON1_ALLMI
35 48 28.4 92 1 GON1_CHICK
36 48 28.4 99 1 GON1_DICLA
37 48 28.4 3133 1 HMCT_BOMMO
38 47.5 28.1 94 1 GON1_HAPBU
39 47.5 28.1 668 1 VNCS_MEVA
40 47.5 28.1 725 1 ADDB_MOUSE
41 47.5 28.1 725 1 ADDB_MOUSE
42 47 27.8 275 1 APAH_HAEIN
43 47 27.8 1032 1 ITA4_XENLA
44 46.5 27.5 110 1 YHBJ_ACTAC
45 46.5 27.5 440 1 FU10_ARATH

ALIGNMENTS

RESULT 1
CSP_PLARE STANDARD; PRT; 388 AA.
AC P26694;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID:5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:91201303; PubMed:2016283;
RA Lal A.A., Goldman I.F.;
RT "Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria parasite evolutionarily related to the human malaria parasite Plasmodium falciparum.";
RL J. Biol. Chem. 266:6886-6889(1991).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant surface antigen on the sporozoite (the infective stage of the malaria parasite that is transmitted from the mosquito to the vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for anchoring the protein to the cell membrane. The repeat sequences would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M60972; AAA29561.1; -
PIR; A39756; A39756.
InterPro; IPR003067; Crcmsprzoite.
InterPro; IPR000884; TSP1.
Pfam; PF00050; TSP_1; 1.
PRINTS; PR01303; Crcmsprzoite.
SMART; SM00209; TSP1; 1.
PROSITE; PSS0092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 388 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 120 263 37 X AA TANDEM REPEATS OF N-[AV]-[ND] - P.
FT REPEAT 120 123 P.
FT REPEAT 124 127 1.
FT REPEAT 128 131 2.
FT REPEAT 132 135 3.
FT REPEAT 136 139 4.
FT REPEAT 140 143 5.
FT REPEAT 144 144 6.
FT REPEAT 144 144 7.


```

FT REPEAT 148 151 8.
FT REPEAT 152 155 9.
FT REPEAT 156 159 10.
FT REPEAT 160 163 11.
FT REPEAT 164 167 12.
FT REPEAT 168 171 13.
FT REPEAT 172 175 14.
FT REPEAT 176 179 15.
FT REPEAT 180 183 16.
FT REPEAT 184 187 17.
FT REPEAT 188 191 18.
FT REPEAT 192 195 19.
FT REPEAT 196 199 20.
FT REPEAT 200 203 21.
FT REPEAT 204 207 22.
FT REPEAT 208 211 23.
FT REPEAT 212 215 24.
FT REPEAT 216 219 25.
FT REPEAT 220 223 26.
FT REPEAT 224 227 27.
FT REPEAT 228 231 28.
FT REPEAT 232 235 29.
FT REPEAT 236 239 30.
FT REPEAT 240 243 31.
FT REPEAT 244 247 32.
FT REPEAT 248 251 33.
FT REPEAT 252 255 34.
FT REPEAT 256 259 35.
FT REPEAT 260 263 36.
FT REPEAT 264 267 37.
FT DOMAIN 313 366 TSP TYPE-1.
SQ SEQUENCE 388 AA; 42245 MW; C031BEFBE2E35604 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 388;
Best Local Similarity 78.3%; Pred. No. 0.00033;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Cy 2 EKKIAKMEKASSVFNVNSGSL 24
Db 356 EKKICKMEKCSVFNVNSISGL 378

RESULT 2
CSP_PLAFO STANDARD; PRT; 397 AA.
ID P19597; Q25798;
AC 01-FEB-1991 (Rel. 17, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345189; PubMed=2668895;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
candidate vaccine antigen.";
RL Nucleic Acids Res. 17:5854-5854 (1989).
RN [2]
RP REVISIONS.
RA Campbell J.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92152598; PubMed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
RA Thomas A.W., Baqar S., Cochran M.A., Thanassi J., Levine M.N.,
RA Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human
infectivity of a cloned line";
RL Exp. Parasitol. 74:159-168 (1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=8936498; PubMed=2671723;
RA Caspers P., Gentz R., Matile H., Pink J.R., Sinigaglia F.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
isolate used in malaria vaccine trials.";
RL Mol. Biochem. Parasitol. 35:185-190 (1989).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
surface antigen on the sporozoite (the infective stage of the
malaria parasite that is transmitted from the mosquito to the
vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for
anchoring the protein to the cell membrane. The repeat sequences
would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X15363; CAA33421.1; -
DR EMBL; M83886; AAA29521.1; -
DR EMBL; M22982; AAA29527.1; -
DR PIR; S05428; S05428.
DR InterPro; IPR003067; Circmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
KW SIGNAL
FT CHAIN 1 16
FT DOMAIN 105 272
FT REPEAT 105 108 1.
FT REPEAT 109 112 2.
FT REPEAT 113 116 3.
FT REPEAT 117 120 4.
FT REPEAT 121 124 5.
FT REPEAT 125 128 6.
FT REPEAT 129 132 7.
FT REPEAT 133 136 8.
FT REPEAT 137 140 9.
FT REPEAT 141 144 10.
FT REPEAT 145 148 11.
FT REPEAT 149 152 12.
FT REPEAT 153 156 13.
FT REPEAT 157 160 14.
FT REPEAT 161 164 15.
FT REPEAT 165 168 16.
FT REPEAT 169 172 17.
FT REPEAT 173 176 18.
FT REPEAT 177 180 19.
FT REPEAT 181 184 20.
FT REPEAT 185 188 21.
FT REPEAT 189 192 22.
FT REPEAT 193 196 23.
FT REPEAT 197 200 24.
FT REPEAT 201 204 25.
FT REPEAT 205 208 26.
FT REPEAT 209 212 27.
FT REPEAT 213 216 28.
FT REPEAT 217 220 29.
FT REPEAT 221 224 30.
FT REPEAT 225 228 31.
FT REPEAT 229 232 32.
FT REPEAT 233 236 33.
FT REPEAT 237 240 34.
FT REPEAT 241 244 35.

```

```

FT REPEAT 245 248 36.
FT REPEAT 249 252 37.
FT REPEAT 253 256 38.
FT REPEAT 257 260 39.
FT REPEAT 261 264 40.
FT REPEAT 265 268 41.
FT REPEAT 269 272 42.
FT DOMAIN 322 375
FT CONFLICT 194 194
SQ SEQUENCE 397 AA; 42646 MW; 9E81146F59EBCEA3 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 397;
Best Local Similarity 78.3%; Pred. No. 0.00034;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKXIARMEKASSVFNVNSGSL 24
Db 365 EKXICKMEKCSSVFNVNSSIGL 387

RESULT 3
CSP_PLAFA
ID - CSP_PLAFA STANDARD; PRT; 412 AA.
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID:5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hochmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy E.P., Diggs C.L., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984)
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for
CC anchoring the protein to the cell membrane. The repeat sequences
CC would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announc/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; K02194; AAA29524.1; -
DR PIR; A03388; OZZCAF.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1.1.
DR PRINTS; PR01303; Crcmsprzoite.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS0092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 286 41 X 4 AA TANDEM REPEATS OF P-N-[AV]-
FT REPEAT 123 126 [ND].
FT REPEAT 127 130 1.
FT REPEAT 131 134 2.
FT REPEAT 133 134 3.

```

```

FT REPEAT 135 138 4.
FT REPEAT 139 142 5.
FT REPEAT 143 146 6.
FT REPEAT 147 150 7.
FT REPEAT 151 154 8.
FT REPEAT 155 158 9.
FT REPEAT 159 162 10.
FT REPEAT 163 166 11.
FT REPEAT 167 170 12.
FT REPEAT 171 174 13.
FT REPEAT 175 178 14.
FT REPEAT 179 182 15.
FT REPEAT 183 186 16.
FT REPEAT 187 190 17.
FT REPEAT 191 194 18.
FT REPEAT 195 198 19.
FT REPEAT 199 202 20.
FT REPEAT 203 206 21.
FT REPEAT 207 210 22.
FT REPEAT 211 214 23.
FT REPEAT 215 218 24.
FT REPEAT 219 222 25.
FT REPEAT 223 226 26.
FT REPEAT 227 230 27.
FT REPEAT 231 234 28.
FT REPEAT 235 238 29.
FT REPEAT 239 242 30.
FT REPEAT 243 246 31.
FT REPEAT 247 250 32.
FT REPEAT 251 254 33.
FT REPEAT 255 258 34.
FT REPEAT 259 262 35.
FT REPEAT 263 266 36.
FT REPEAT 267 270 37.
FT REPEAT 271 274 38.
FT REPEAT 275 278 39.
FT REPEAT 279 282 40.
FT REPEAT 283 286 41.
FT DOMAIN 337 390 TSP TYPE-1.
SQ SEQUENCE 412 AA; 44420 MW; 1E5EED3DE9065F8 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 412;
Best Local Similarity 78.3%; Pred. No. 0.00036;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKXIARMEKASSVFNVNSGSL 24
Db 380 EKXICKMEKCSSVFNVNSSIGL 402

RESULT 4
CSP_PLAFA
ID - CSP_PLAFA STANDARD; PRT; 424 AA.
AC P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID:5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Znea V.;
RT "Circumsporozoite gene of a Plasmodium falciparum strain from
RT Thailand.";
RL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for

```

CC anchoring the protein to the cell membrane. The repeat sequences
CC would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19752; AAA29555.1; --
DR PIR; A54533; A54533.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF000090; TSP1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
DR Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 424
FT DOMAIN 123 298
FT REPEAT 123 126
FT REPEAT 127 130
FT REPEAT 131 134
FT REPEAT 135 138
FT REPEAT 139 142
FT REPEAT 143 146
FT REPEAT 147 150
FT REPEAT 151 154
FT REPEAT 155 158
FT REPEAT 159 162
FT REPEAT 163 166
FT REPEAT 167 170
FT REPEAT 171 174
FT REPEAT 175 178
FT REPEAT 179 182
FT REPEAT 183 186
FT REPEAT 187 190
FT REPEAT 191 194
FT REPEAT 195 198
FT REPEAT 199 202
FT REPEAT 203 206
FT REPEAT 207 210
FT REPEAT 211 214
FT REPEAT 215 218
FT REPEAT 219 222
FT REPEAT 223 226
FT REPEAT 227 230
FT REPEAT 231 234
FT REPEAT 235 238
FT REPEAT 239 242
FT REPEAT 243 246
FT REPEAT 247 250
FT REPEAT 251 254
FT REPEAT 255 258
FT REPEAT 259 262
FT REPEAT 263 266
FT REPEAT 267 270
FT REPEAT 271 274
FT REPEAT 275 278
FT REPEAT 279 282
FT REPEAT 283 286
FT REPEAT 287 290
FT REPEAT 291 294
FT REPEAT 295 298
FT DOMAIN 349 402
SQ SEQUENCE 424 AA; 45610 MW; 710AB14238786CD9 CRC64;
Query Match 48.5%; Score 82; DB 1; Length 424;

Best Local Similarity 78.3%; Pred. No. 0.00037;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 EKIKAKMEKASSVFNVNSSIGL 24
DB 392 EKIKAKMEKASSVFNVNSSIGL 414
RESULT 5
CSP_FLAFW STANDARD; PRT; 442 AA.
ID CSP_FLAFW
AC P08307;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115616; PubMed=3543671;
RA Lockyer M.J.; Schwarz R.T.;
RT "Strain variation in the circumsporozoite protein gene of Plasmodium
RT falciparum.";
RL Mol. Biochem. Parasitol. 22:101-108(1987).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for
CC anchoring the protein to the cell membrane. The repeat sequences
CC would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15505; AAA29554.1; --
DR PIR; A54529; A54529.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF000090; TSP1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 442
FT DOMAIN 134 317
FT REPEAT 134 137
FT REPEAT 138 141
FT REPEAT 142 145
FT REPEAT 146 149
FT REPEAT 150 153
FT REPEAT 154 157
FT REPEAT 158 161
FT REPEAT 162 165
FT REPEAT 166 169
FT REPEAT 170 173
FT REPEAT 174 177
FT REPEAT 178 181
FT REPEAT 182 185
FT REPEAT 186 189
FT REPEAT 190 193
FT REPEAT 194 197
FT REPEAT 198 201
FT REPEAT 201 201
CIRCUMSPOROZOITE PROTEIN.
46 X 4 AA TANDEM REPEATS OF N-[AV] - [ND] -
P.

```

FT REPEAT 202 205 18.
FT REPEAT 206 209 19.
FT REPEAT 210 213 20.
FT REPEAT 214 217 21.
FT REPEAT 218 221 22.
FT REPEAT 222 225 23.
FT REPEAT 226 229 24.
FT REPEAT 230 233 25.
FT REPEAT 234 237 26.
FT REPEAT 238 241 27.
FT REPEAT 242 245 28.
FT REPEAT 246 249 29.
FT REPEAT 250 253 30.
FT REPEAT 254 257 31.
FT REPEAT 258 261 32.
FT REPEAT 262 265 33.
FT REPEAT 266 269 34.
FT REPEAT 270 273 35.
FT REPEAT 274 277 36.
FT REPEAT 278 281 37.
FT REPEAT 282 285 38.
FT REPEAT 286 289 39.
FT REPEAT 290 293 40.
FT REPEAT 294 297 41.
FT REPEAT 298 301 42.
FT REPEAT 302 305 43.
FT REPEAT 306 309 44.
FT REPEAT 310 313 45.
FT REPEAT 314 317 46.
FT DOMAIN 367 420 TSP TYPE-1.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 48.5%; Score 82; DB 1; Length 442;
Best Local Similarity 78.3%; Pred. No. 0.00038;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKXIAKMEKASVFNVSGL 24
DB 410 EKXICKMEKCSVFNVSIGL 432

RESULT 6
GONI_RAT
ID GONI_RAT STANDARD; PRT; 92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GNRH I) (Luliberin I); Prolactin release-inhibiting factor
DE I].
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression.";
RL Mol. Endocrinol. 3:1257-1262(1989).
RN (3)

SEQUENCE FROM N.A.
TISSUE=Thymus;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN (4)
RP SEQUENCE OF 1-47 FROM N.A.
TISSUE=Heart;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus";
RL Science 235:1514-1517(1987).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Central nervous system.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -
DR EMBL; M12579; AAA41263.1; -
DR EMBL; M31670; AAA41264.1; -
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -
DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
DR PIR; A40147; RERTG.
DR InterPro; IPR002012; GNRH.
DR InterPro; IPR004079; GonadolibereinI.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match 33.4%; Score 56.5; DB 1; Length 92;
Best Local Similarity 48.3%; Pred. No. 0.35;
Matches 14; Conservative 2; Mismatches 12; Indels 1; Gaps 1;

QY 5 IAKMEKASVFNVSGL 32
DB 4 IPKMAAVVLLTVCLGCSQHSYGLRP 32

RESULT 7
CSP_PLABE STANDARD; PRT; 339 AA.
AC P06915;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
DE Plasmodium berghei.
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

```


FT REPEAT 216 217 2-2.
 FT REPEAT 218 219 2-3 (APPROXIMATE).
 FT REPEAT 220 221 2-4.
 FT REPEAT 222 223 2-5.
 FT REPEAT 224 225 2-6.
 FT REPEAT 226 227 2-7.
 FT REPEAT 228 229 2-8.
 FT REPEAT 230 231 2-9.
 FT REPEAT 232 233 2-10.
 FT REPEAT 234 235 2-11.
 FT REPEAT 236 237 2-12.
 FT REPEAT 238 239 2-13 (APPROXIMATE).
 FT REPEAT 240 241 2-14.
 FT REPEAT 242 243 2-15.
 FT REPEAT 244 245 2-16.
 FT REPEAT 246 247 2-17 (APPROXIMATE).
 FT DOMAIN 274 325 TSP TYPE-1.
 SQ SEQUENCE 347 AA; 37776 MW; 0EC240EE35681AF8 CRC64;

Query Match 32.5%; Score 55; DB 1; Length 347;
 Best Local Similarity 47.4%; Pred.No. 2.3;
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EKXIADKXKSSIFNVNVS 20
 Db 315 DTEICKMDKXSSIFNVN 333

RESULT 9

CSP PLAYO STANDARD; PRT; 367 AA.
 ID P06914;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium berghei yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87137555; PubMed=3102479;
 RA Lal A.A., de la Cruz V.F., Welsh J.A., Charoenvit Y., Maloy W.L.,
 RA McCutchan T.F.;
 RT "Structure of the gene encoding the circumsporozoite protein of
 RT Plasmodium yoelii. A rodent model for examining antimalarial
 RT sporozoite vaccines";
 RL J. Biol. Chem. 262:2937-2940(1987).
 RN [2]
 RP SEQUENCE OF 1-140 AND 260-367 FROM N.A.
 RX MEDLINE=88232798; PubMed=3287156;
 RA de la Cruz V.F., Lal A.A., McCutchan T.F.;
 RT "Variation among circumsporozoite protein genes from rodent
 RT malariae";
 RL Mol. Biochem. Parasitol. 28:31-38(1988).
 CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
 CC surface antigen on the sporozoite (the infective stage of the
 CC malaria parasite that is transmitted from the mosquito to the
 CC vertebrate host).
 CC -!- MISCELLANEOUS: The C-terminal region is probably used for
 CC anchoring the protein to the cell membrane. The repeat sequences
 CC would be the surface antigen of the organism.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J02695; AAA29556.1; -.

DR EMBL; M18821; AAA29559.1; -.
 DR EMBL; M22698; AAA29560.1; -.
 DR InterPro; IPR003067; Circmsporzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1
 DR FRINIS; PRO1303; CIRCMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS0092; TSP1; 1.
 DR Malaria; Sporozite; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 367 CIRCUMSPOROZOITE PROTEIN.
 FT DOMAIN 139 228 15 X 6 AA TANDEM REPEATS OF Q-G-P-G-A-P.
 FT REPEAT 139 144 1-1.
 FT REPEAT 145 150 1-2.
 FT REPEAT 151 156 1-3.
 FT REPEAT 157 162 1-4.
 FT REPEAT 163 168 1-5.
 FT REPEAT 169 174 1-6.
 FT REPEAT 175 180 1-7.
 FT REPEAT 181 186 1-8.
 FT REPEAT 187 192 1-9.
 FT REPEAT 193 198 1-10.
 FT REPEAT 199 204 1-11.
 FT REPEAT 205 210 1-12.
 FT REPEAT 211 216 1-13.
 FT REPEAT 217 222 1-14.
 FT REPEAT 223 228 1-15.
 FT DOMAIN 229 260 8 X 4 AA APPROXIMATE TANDEM REPEATS OF Q-P-P.
 FT REPEAT 229 232 2-1 (APPROXIMATE).
 FT REPEAT 233 236 2-2.
 FT REPEAT 237 240 2-3.
 FT REPEAT 241 244 2-4.
 FT REPEAT 245 248 2-5.
 FT REPEAT 249 252 2-6.
 FT REPEAT 253 256 2-7.
 FT REPEAT 257 260 2-8 (APPROXIMATE).
 FT DOMAIN 294 345 TSP TYPE-1.
 SQ SEQUENCE 367 AA; 38888 MW; 1EA56AFF7FFCB5E3 CRC64;

Query Match 32.5%; Score 55; DB 1; Length 367;
 Best Local Similarity 47.4%; Pred.No. 2.5;
 Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EKXIADKXKSSIFNVNVS 20
 Db 335 DTEICKMDKXSSIFNVN 353

RESULT 10

GONI MACMU STANDARD; PRT; 67 AA.
 ID GONI MACMU
 AC P55247;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadolibirin I precursor [Contains: Gonadolibirin I (LH-RH I)
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (luliberin I); GnRH-associated peptide I]
 DE (Fragment).
 GN GNRH1 OR GNRH OR LHRH.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hypothalamus;
 RX MEDLINE=95124501; PubMed=7545971;
 RA Ma Y.J., Costa M.E., Ojeda S.R.;
 RT "Developmental expression of the genes encoding transforming growth
 RT factor alpha and its receptor in the hypothalamus of female rhesus

RT macaques.";
 RL Neuroendocrinology 60:346-359(1994).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch).
 CC
 CC EMBL; M75918; AAB33096.1; --
 DR PIR; I78541; I78541; GnRH.
 DR InterPro; IPR002012; GnRH.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERNI.
 DR PROSITE; PS00473; GnRH; 1.
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Pyroglutamate carboxylic acid.
 FT NON_TER 1 1
 FT SIGNAL <1 5 BY SIMILARITY.
 FT CHAIN 6 >67 PROGONADOLIBERIN I.
 FT PEPTIDE 6 15 GONADOLIBERIN I.
 FT PEPTIDE 19 >67 GnRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY (BY SIMILARITY).
 FT MOD_RES 6 6 PYROLIDONE CARBOXYLIC ACID (BY
 FT SIMILARITY).
 FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
 FT SIMILARITY).
 FT NON_TER 67 67
 FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
 SQ
 Query Match 32.0%; Score 54; DB 1; Length 67;
 Best Local Similarity 90.0%; Pred. No. 0.57; 1; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 23 SLHWSYGLRP 32
 Db 5 SQHWSYGLRP 14
 | | | | | | | | | |
 RESULT 11
 GONI_MOUSE STANDARD; PRT; 90 AA.
 ID GONI_MOUSE
 AC P13562;
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
 DE I].
 GN GnRH1 OR GnRH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87069928; PubMed=3024317;
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolics K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 RT responsible for hypogonadism in the hpg mouse.";
 RL Science 234:1366-1371(1986).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates

CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch).
 CC
 CC EMBL; M14872; AAA37717.1; --
 DR PIR; A47578; REMSG.
 DR MGD; MGI:95789; GnRH.
 DR InterPro; IPR002012; GnRH.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERNI.
 DR PROSITE; PS00473; GnRH; 1.
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal; Pyroglutamate carboxylic acid.
 FT SIGNAL 1 21
 FT CHAIN 22 90 PROGONADOLIBERIN I.
 FT PEPTIDE 22 31 GONADOLIBERIN I.
 FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
 FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 22 22 PYROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
 FT SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;
 SQ
 Query Match 32.0%; Score 54; DB 1; Length 90;
 Best Local Similarity 90.0%; Pred. No. 0.78; 1; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 23 SLHWSYGLRP 32
 Db 21 SQHWSYGLRP 30
 | | | | | | | | | |
 RESULT 12
 GONI_FIG STANDARD; PRT; 91 AA.
 ID GONI_FIG
 AC P49921; Created
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].
 GN GnRH1 OR GnRH.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Hypothalamus;
 RA Weesner G.D., Matteri R.L., Becker B.A.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=72114303; PubMed=4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses.";
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN
 RX MEDLINE=72065376; PubMed=4942726;

RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RL phase method."; Biochem. Biophys. Res. Commun. 45:822-827(1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RA MEDLINE=72117544; PubMed=4946275;
 RX Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L32864; AAA31066.1; -.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadoliberin.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBRN.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal; Pyroglutamate carboxylic acid.
 FT SIGNAL 1 23
 FT CHAIN 24 91 PROGONADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 34 91 GnRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 24 24 PYROGLUTAMATE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
 Query Match 32.0%; Score 54; DB 1; Length 91;
 Best Local Similarity 90.0%; Pred. No. 0.79;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 23 SLHWSYGLRP 32 PRT; 92 AA.
 Db 23 SCHWSYGLRP 32
 | | | | | | | | | |
 RESULT 13
 GONL_HUMAN STANDARD; PRT; 92 AA.
 AC P01148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (luliberin I) (Gonadorelin); GnRH-associated
 DE peptide I].
 DE GNRH1 OR GNRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8936682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene.";

RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT SER-16.
 RX MEDLINE=85012739; PubMed=6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 RT releasing hormone";
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide
 RT biosynthesized in the human placenta";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 RN [5]
 RP VARIANT SER-16.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 RT of human genes";
 RL Nat. Genet. 22:231-238(1999).
 RN [6]
 RP ERRATUM.
 RX Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 RA Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACOLOGICAL: Available under the names Factrel (Ayerst Labs),
 CC Lutrepulse or Lutrelaf (Ferring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X01059; CAA25526.1; -.
 DR EMBL; M12578; AAA35916.1; -.
 DR EMBL; X15215; CAA33285.1; -.
 DR PIR; S05308; RHUG.
 DR Genew; HGNC:4419; GNRH1.
 DR MIN; 152760; -.
 DR GO; GO:0005625; C:soluble fraction; TAS.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadoliberin.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBRN.

FT MOD RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
Query Match 30.8%; Score 52; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 HWSYGLRP 32
Db 2 HWSYGLRP 9

Search completed: March 10, 2004, 09:13:55
Job time : 7.03502 Secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein : protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 10.8405 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFIGITELGSLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B COMB.pep: *
5: /cgn2_6/prodata/2/1aa/FACTUS COMB.pep: *
6: /cgn2_6/prodata/2/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|--------------------|
| 1 | 119.5 | 80.7 | 47 | 1 | US-08-446-692-35 |
| 2 | 119.5 | 80.7 | 47 | 2 | US-08-488-351A-35 |
| 3 | 112.5 | 76.0 | 27 | 1 | US-08-446-692-13 |
| 4 | 112.5 | 76.0 | 27 | 2 | US-08-488-351A-13 |
| 5 | 81.5 | 55.1 | 31 | 4 | US-09-362-731A-1 |
| 6 | 80 | 54.1 | 32 | 4 | US-09-362-731A-5 |
| 7 | 80 | 54.1 | 37 | 1 | US-08-446-692-57 |
| 8 | 80 | 54.1 | 37 | 1 | US-08-446-692-63 |
| 9 | 80 | 54.1 | 37 | 2 | US-08-488-351A-57 |
| 10 | 80 | 54.1 | 37 | 2 | US-08-488-351A-63 |
| 11 | 80 | 54.1 | 137 | 4 | US-09-362-731A-3 |
| 12 | 78 | 52.7 | 50 | 4 | US-08-945-289-8 |
| 13 | 78 | 52.7 | 182 | 4 | US-09-396-937-16 |
| 14 | 77 | 52.0 | 24 | 4 | US-08-464-496-25 |
| 15 | 77 | 52.0 | 24 | 4 | US-08-197-484-110 |
| 16 | 77 | 52.0 | 24 | 5 | PCT-US92-07218-25 |
| 17 | 77 | 52.0 | 24 | 5 | PCT-US95-02121-110 |
| 18 | 77 | 52.0 | 27 | 4 | US-08-464-496-32 |
| 19 | 77 | 52.0 | 27 | 5 | PCT-US92-07218-32 |
| 20 | 76 | 51.4 | 31 | 4 | US-08-432-483A-2 |
| 21 | 76 | 51.4 | 31 | 4 | US-08-945-289-2 |
| 22 | 74 | 50.0 | 15 | 2 | US-08-319-704-10 |
| 23 | 74 | 50.0 | 15 | 2 | US-08-661-052-6 |
| 24 | 74 | 50.0 | 15 | 2 | US-08-460-502-7 |
| 25 | 74 | 50.0 | 15 | 3 | US-09-046-373-2 |
| 26 | 74 | 50.0 | 15 | 3 | US-09-188-082-6 |
| 27 | 74 | 50.0 | 15 | 4 | US-09-364-088-6 |

Sequence 6, Appli
Sequence 7, Appli
Sequence 4, Appli
Sequence 13, Appli
Sequence 34, Appli
Sequence 1, Appli
Sequence 69, Appli
Sequence 55, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 20, Appli
Sequence 23, Appli
Sequence 7, Appli
Sequence 63, Appli
Sequence 18, Appli
Sequence 17, Appli
Sequence 16, Appli

US-09-102-716-6
US-09-148-711A-7
US-08-716-249-4
US-09-362-731A-13
US-09-396-937-34
US-09-405-988A-1
PCT-US93-11703-69
US-09-248-588-55
US-08-446-692-4
US-08-488-351A-4
US-09-100-409A-40
US-08-472-701-23
PCT-US95-08596-23
PCT-US95-13841-7
PCT-US93-11703-63
US-09-396-937-18
US-08-913-880C-17
US-08-913-880C-16

ALIGNMENTS

RESULT 1

US-08-446-692-35
; Sequence 35 Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-35

Query Match 80.7%; Score 119.5; DB 1; Length 47;
Best Local Similarity 88.9%; Pred. No. 8e-12;
Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 QYIKANSKFIGITELGSLHWSYGLRP 27

DB 21 QYIKANSKFIGITELGGE-HWSYGLRP 46

RESULT 2
US-08-488-351A-35
; Sequence 35, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-35
Query Match 80.7%; Score 119.5; DB 2; Length 47;
Best Local Similarity 88.9%; Pred. No. 8e-12; 2; Indels 1; Gaps 1;
Matches 24; Conservative 0; Mismatches 1; Gaps 1;
QY 1 QYKANSKFIGITELGSLHWSYGLRP 27
Db 21 QYKANSKFIGITELGSG-HWSYGLRP 46
RESULT 3
US-08-446-692-13
; Sequence 13, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-35

; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-13
Query Match 76.0%; Score 112.5; DB 1; Length 27;
Best Local Similarity 85.2%; Pred. No. 5.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 QYKANSKFIGITELGSLHWSYGLRP 27
Db 3 QYKANSKFIGITELG---HWSYGLRP 26
RESULT 4
US-08-488-351A-13
; Sequence 13, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-13

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-13

Query Match 76.0%; Score 112.5; DB 2; Length 27;
Best Local Similarity 85.2%; Pred. No. 5.4e-11;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 QYKANSKFIGITELGPSLHWSYGLRP 27
Db 3 QYKANSKFIGITELE---HWSYGLRP 26

RESULT 5
US-09-362-731A-1
; Sequence 1, Application US/09362731A
; Patent No. 6602509
; GENERAL INFORMATION:
; APPLICANT: SAINT-REMY, Jean-Marie
; APPLICANT: JACQUEMIN, Marc
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF AL
; FILE REFERENCE: 99-0720*/LC/01699
; CURRENT APPLICATION NUMBER: US/09/362,731A
; CURRENT FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Version 2.1
; SEQ ID NO 1
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; US-09-362-731A-1

Query Match 55.1%; Score 81.5; DB 4; Length 31;
Best Local Similarity 61.3%; Pred. No. 4.9e-06;
Matches 19; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY 1 QYKANSKFIGITELG-----PSLHWS 22
Db 1 QYKANSKFIGITELGGHKEIKVLPGCHGS 31

RESULT 6
US-09-362-731A-5
; Sequence 5, Application US/09362731A
; Patent No. 6602509
; GENERAL INFORMATION:
; APPLICANT: SAINT-REMY, Jean-Marie
; APPLICANT: JACQUEMIN, Marc
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF AL
; FILE REFERENCE: 99-0720*/LC/01699
; CURRENT APPLICATION NUMBER: US/09/362,731A
; CURRENT FILING DATE: 1999-07-29

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Version 2.1
; SEQ ID NO 5
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; US-09-362-731A-5

Query Match 54.1%; Score 80; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELG 16
Db 1 QYKANSKFIGITELG 16

RESULT 7
US-08-446-692-57
; Sequence 57, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-57

Query Match 54.1%; Score 80; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELG 16
Db 21 QYKANSKFIGITELG 36

RESULT 8
US-08-446-692-63
; Sequence 63, Application US/08446692
```

```

; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; US-08-446-692-63

Query Match 54.1%; Score 80; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
Db 5 QYIKANSKFIGITELG 20

RESULT 9
US-08-488-351A-57
; Sequence 57 Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A

```

```

; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-57

Query Match 54.1%; Score 80; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
Db 21 QYIKANSKFIGITELG 36

RESULT 10
US-08-488-351A-63
; Sequence 63 Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424

```

```

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-63

Query Match 54.1%; Score 80; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELG 16
Db 5 QYKANSKFIGITELG 20

RESULT 11
US-09-362-731A-3
; Sequence 3, Application US/09362731A
; Patent No. 6602509
; GENERAL INFORMATION:
; APPLICANT: SAINT-REMY, Jean-Marie
; APPLICANT: JACQUEMIN, Marc
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF AL
; FILE REFERENCE: 99-0720*/LC/01699
; CURRENT APPLICATION NUMBER: US/09/362,731A
; CURRENT FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 3
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
; US-09-362-731A-3

Query Match 54.1%; Score 80; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELG 16
Db 2 QYKANSKFIGITELG 17

RESULT 12
US-08-945-289-8
; Sequence 8, Application US/08945289
; Patent No. 6555113
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles, W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER
; TITLE OF INVENTION: TRANSFER PROTEIN (CETP) ACTIVITY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yankwich & Associates
; STREET: 130 Bishop Allen Drive
; CITY: Cambridge
; STATE: Massachusetts

;
; COUNTRY: USA
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,289
; FILING DATE: October 17, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/432,483
; FILING DATE: May 1, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Leon R. Yankwich
; REGISTRATION NUMBER: 30,237
; REFERENCE/DOCKET NUMBER: TCS-411.1P US
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL:
; ANTI-SENSE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; US-08-945-289-8

Query Match 52.7%; Score 78; DB 4; Length 50;
Best Local Similarity 94.1%; Pred. No. 3e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITELGP 17
Db 2 QYKANSKFIGITELFP 18

RESULT 13
US-09-396-937-16
; Sequence 16, Application US/09396937
; Patent No. 6645500
; GENERAL INFORMATION:
; APPLICANT: M&E Biotech A/S
; APPLICANT: HARKIER, Torben
; APPLICANT: HANING, Jesper
; TITLE OF INVENTION: Method for Down-Regulating Osteoprotegerin Ligand
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 22021 PC 1
; CURRENT APPLICATION NUMBER: US/09/396,937
; CURRENT FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion
; OTHER INFORMATION: between murine OPGL residues 158-316 with tetanus
; OTHER INFORMATION: toxoid P2 epitope introduced, and His tag
; US-09-396-937-16

Query Match 52.7%; Score 78; DB 4; Length 182;
Best Local Similarity 77.3%; Pred. No. 0.00013;
Matches 17; Conservative 1; Mismatches 0; Indels 4; Gaps 1;

QY 1 QYKANSKFIGITELGPSLHWS 22
Db 113 QYKANSKFIGITE----LNWS 130
```

```
RESULT 14
US-08-464-496-25
; Sequence 25, Application US/08464496
; Patent No. 632789
; GENERAL INFORMATION:
; APPLICANT: Vitello, Maria
; APPLICANT: Chesnut, Robert
; TITLE OF INVENTION: HLA-RESTRICTED HEPATITIS B VIRUS CTL
; TITLE OF INVENTION: EPITOPE
; FILE REFERENCE: 39963-20001.13
; CURRENT APPLICATION NUMBER: US/08/464,496
; CURRENT FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 07/935,811
; PRIOR FILING DATE: 1992-08-26
; PRIOR APPLICATION NUMBER: 07/874,491
; PRIOR FILING DATE: 1992-04-27
; PRIOR APPLICATION NUMBER: 07/827,682
; PRIOR FILING DATE: 1992-01-29
; PRIOR APPLICATION NUMBER: 07/749,568
; PRIOR FILING DATE: 1991-08-26
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Oligopeptide
US-08-464-496-25

Query Match 52.0%; Score 77; DB 4; Length 24;
Best Local Similarity 88.9%; Pred. No. 1.9e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFGITELGPS 18
Db 1 QYIKANSKFGITELGPS 18

RESULT 15
US-08-197-484-110
; Sequence 110, Application US/08197484
; Patent No. 8419331
; GENERAL INFORMATION:
; APPLICANT: VITIELLO, Maria A.
; APPLICANT: CHESTNUT, Robert W.
; APPLICANT: SETTE, Alessandro D.
; APPLICANT: CELIS, Esteban
; APPLICANT: GRAY, Howard
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-197-484-110

Query Match 52.0%; Score 77; DB 4; Length 24;
Best Local Similarity 88.9%; Pred. No. 1.9e-05;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QYIKANSKFGITELGPS 18
Db 1 QYIKANSKFGITELGPS 18

Search completed: March 10, 2004, 09:28:54
Job time : 10.8405 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: March 10, 2004, 08:58:54 ; Search time 10.2724 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-12
Perfect score: 169
Sequence: 1 DEKIAKMEKASSVFNVNGPSLHWSYGLRXP 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: Pirl: *
2: Pirl2: *
3: Pirl3: *
4: Pirl4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 82 | 48.5 | 388 | 2 A39756 | circumsporozoite p |
| 2 | 82 | 48.5 | 405 | 2 S05428 | circumsporozoite p |
| 3 | 82 | 48.5 | 412 | 1 OZQAP | circumsporozoite p |
| 4 | 82 | 48.5 | 424 | 2 A54533 | circumsporozoite p |
| 5 | 82 | 48.5 | 442 | 2 A54529 | circumsporozoite p |
| 6 | 56.5 | 33.4 | 92 | 1 RHRTG | gonadoliberin prec |
| 7 | 55 | 32.5 | 264 | 2 A44969 | circumsporozoite p |
| 8 | 55 | 32.5 | 332 | 1 OZQMB | circumsporozoite p |
| 9 | 55 | 32.5 | 348 | 1 OZQBK | circumsporozoite p |
| 10 | 55 | 32.5 | 360 | 2 AD2528 | circumsporozoite p |
| 11 | 55 | 32.5 | 367 | 1 OZQMY | circumsporozoite p |
| 12 | 54 | 32.0 | 67 | 2 I78541 | gonadoliberin prec |
| 13 | 54 | 32.0 | 90 | 1 RHMSG | gonadoliberin prec |
| 14 | 54 | 32.0 | 92 | 1 RHUG | gonadoliberin - pi |
| 15 | 52 | 30.8 | 10 | 1 RHGG | gonadoliberin - sh |
| 16 | 52 | 30.8 | 10 | 1 RHSG | gonadoliberin - sh |
| 17 | 52 | 30.8 | 89 | 2 I5423 | gonadoliberin prote |
| 18 | 52 | 30.8 | 994 | 2 S00960 | hypothetical prote |
| 19 | 51 | 30.2 | 91 | 2 JC7393 | medaka-type gonado |
| 20 | 51 | 30.2 | 429 | 2 A54504 | circumsporozoite p |
| 21 | 51 | 30.2 | 485 | 2 A60610 | circumsporozoite p |
| 22 | 51 | 30.2 | 507 | 2 H82378 | probable long-chain |
| 23 | 50 | 29.6 | 1484 | 2 C97196 | probable membrane |
| 24 | 50 | 29.6 | 2339 | 2 A45597 | DNA-directed RNA p |
| 25 | 49.5 | 29.3 | 630 | 2 A36359 | VSG expression sit |
| 26 | 49.5 | 29.3 | 630 | 2 S13724 | ESAG 8 protein - T |
| 27 | 49.5 | 29.3 | 999 | 2 S15961 | hypothetical protei |
| 28 | 49.5 | 29.3 | 1265 | 2 S57968 | ran-binding protei |
| 29 | 49.5 | 29.3 | 3176 | 2 CGH3A | collagen alpha 3(V |

30 49 29.0 532 2 S27373
31 49 29.0 1146 2 S40311
32 48.5 28.7 630 1 BWUT8Q
33 48 28.4 10 1 RHAQ1
34 48 28.4 66 2 T43096
35 48 28.4 92 2 I50644
36 48 28.4 260 2 T41554
37 48 28.4 525 2 T40088
38 48 28.4 3133 2 S52093
39 47.5 28.1 98 2 I50739
40 47.5 28.1 562 2 JQ1036
41 47.5 28.1 664 2 AB1430
42 47.5 28.1 664 2 AH1803
43 47.5 28.1 668 1 UYPVME
44 47.5 28.1 2244 2 F90563
45 47 27.8 217 2 A98196

beta-fructofuranos
integrin - fruit f
regulatory protein
gonadoliberin I -
hypothetical prote
gonadoliberin I pr
hypothetical prote
RHGEF domain cont
hemocytin - silkwo
gonadotropin-relea
63k aducin-like p
hypothetical prote
hypothetical prote
noncapsid protein
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

A39756
Circumsporozoite protein - Plasmodium reichenowi
C:Species: Plasmodium reichenowi
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999
C:Accession: A39756
R/Lal, A.A.; Goldman, I.F.
J. Biol. Chem. 266, 6686-6689, 1991
A:Title: Circumsporozoite protein Gene from Plasmodium reichenowi, a chimpanzee malar
A:Reference number: A39756; MUID:91201303; PMID:2016283
A:Accession: A39756
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-388 <LAL>
A:Cross-references: GB:M60972; NID:G160228; PIDN:AAA29561.1; PID:G160229
C:Superfamily: Circumsporozoite protein; thrombospondin type 1 repeat homology
F:312-366/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 48.5%; Score 82; DB 2; Length 388;
Best Local Similarity 78.3%; Pred. NO. 0.00083;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 EKXIAKMEKASSVFNVNGPSL 24

Db 356 EKXIAKMEKASSVFNVNGPSL 378

RESULT 2

S05428
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)
C:Species: Plasmodium falciparum
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000
C:Accession: S05428; A45527; I60657
R/Campbell, J.R.
Nucleic Acids Res. 17, 5854, 1989
A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate v
A:Reference number: S05428; MUID:89345189; PMID:2568895
A:Accession: S05428
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-405 <CAM>
A:Cross-references: EMBL:X15363
R/Caspers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.
Mol. Biochem. Parasitol. 35, 185-190, 1989
A:Title: The circumsporozoite protein Gene from NF54, a Plasmodium falciparum isolate
A:Reference number: A45527; MUID:89364998; PMID:2671723
A:Accession: A45527
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-405 <CAS>
A:Cross-references: GB:M22982; GB:J04650; NID:G160168; PIDN:AAA29527.1; PID:G160169
R/Lockyer, M.J.; Marsh, K.; Newbold, C.I.

Mol. Biochem. Parasitol. 37, 275-280, 1999
A;Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell epitopes
A;Reference number: A60657; MUID:90114334; PMID:2481827
A;Accession: I60657
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 319-336,354-373 <LOC>
C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F;329-383/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 48.5%; Score 82; DB 2; Length 405;
Best Local Similarity 78.3%; Pred. No. 0.00087;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNSGSL 24
|||||
DB 373 EKKICKMEKCSSVFNVNNSIGL 395
|||||

RESULT 3
OZZQAF
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate IMTM22)
C;Species: Plasmodium falciparum
C;Date: 15-Nov-1994 #sequence_revision 15-Nov-1994 #text_change 09-Jun-2000
C;Accession: A03388
R;Dane, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
Science 225, 593-599, 1984
A;Title: Structure of the gene encoding the immunodominant surface antigen on the sporozoite
A;Reference number: A03388; MUID:84250215; PMID:6204383
A;Accession: A03388
A;Molecule type: DNA
A;Residues: 1-412 <DAM>
A;Cross-references: GB:K02194; NID:G160160; PIDN:AAA29524.1; PID:G160161
A;Experimental source: clone 768
C;Comment: Residues 1-16 are the probable signal sequence.
C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F;336-390/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 48.5%; Score 82; DB 1; Length 412;
Best Local Similarity 78.3%; Pred. No. 0.00089;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNSGSL 24
|||||
DB 380 EKKICKMEKCSSVFNVNNSIGL 402
|||||

RESULT 4
A54533
Circumsporozoite protein - malaria parasite (Plasmodium falciparum) (strain T4, Thailand)
C;Species: Plasmodium falciparum
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C;Accession: A54533
R;del Portillo, H.A.; Nuesenztweig, R.S.; Enea, V.
Mol. Biochem. Parasitol. 24, 289-294, 1987
A;Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
A;Reference number: A54533; MUID:87315205; PMID:3306373
A;Accession: A54533
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-424
A;Cross-references: GB:MI9752; NID:G160216; PIDN:AAA29555.1; PID:G160217
C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F;346-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 48.5%; Score 82; DB 2; Length 424;
Best Local Similarity 78.3%; Pred. No. 0.00092;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNSGSL 24
|||||
DB 392 EKKICKMEKCSSVFNVNNSIGL 414
|||||

Mol. Biochem. Parasitol. 37, 275-280, 1999
A;Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell epitopes
A;Reference number: A60657; MUID:90114334; PMID:2481827
A;Accession: I60657
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 319-336,354-373 <LOC>
C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F;329-383/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 48.5%; Score 82; DB 2; Length 442;
Best Local Similarity 78.3%; Pred. No. 0.00096;
Matches 18; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKKIAKMEKASSVFNVNNSGSL 24
|||||
DB 410 EKKICKMEKCSSVFNVNNSIGL 432
|||||

RESULT 6
RHTGT
Gonadoliberin precursor - rat
A;Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormone N-terminus; gonadoliberin; prolactin release-inhibiting factor
N;Contains: gonadoliberin; prolactin release-inhibiting factor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C;Accession: A40147; B26173; A48410
R;Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A;Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic e
A;Reference number: A40147; MUID:89384661; PMID:2476669
A;Accession: A40147
A;Molecule type: DNA
A;Residues: 1-92 <BON>
A;Cross-references: GB:M31670; NID:G204447; PIDN:AAA41264.1; PID:G204448
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonad
A;Reference number: A34090; MUID:86094338; PMID:2867548
A;Accession: B26173
A;Molecule type: mRNA
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12579; NID:G204445; PIDN:AAA41263.1; PID:G204446
R;Maier, C.C.; Marchetti, B.; LeBoeuff, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A;Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormo
A;Reference number: A48410; MUID:93105480; PMID:1468115
A;Accession: A48410
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-92 <MAL>
A;Cross-references: GB:S50870; NID:G262059; PIDN:ARB24572.1; PID:G262060
A;Experimental source: thymus
A;Note: sequence extracted from NCBI backbone (NCBI:121082, NCBI:121083)
C;Genetics:
A;Introns: 47/3; 79/3
C;Function:
A;Description: stimulates pituitary secretion of lutropin and follitropin
A;Note: Gonadoliberin-associated protein may have prolactin release inhibiting activity
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid;
F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-92/Product: progonaoliberin #status predicted <PON>
 F;24-33/Product: gonaoliberin #status predicted <GIN>
 F;37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
 F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 33.4%; Score 56.5; DB 1; Length 92;
 Best Local Similarity 48.3%; Pred. No. 0.72; Indels 1; Gaps 1;
 Matches 14; Conservative 2; Mismatches 12

QY 5 IAKYKAKASSVFNVNSG-PSLHWSYGLRP 32
 Db 4 IPKLMVAIVLLTVLCGSCSOHNSYGLRP 32

RESULT 7
 A44969
 Circumsporozoite protein precursor - Plasmodium yoelii nigeriensis
 C;Species: Plasmodium yoelii nigeriensis
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-1995
 C;Accession: A44969
 R;Colomer-Gould, V.; Enea, V.
 Mol. Biochem. Parasitol. 43, 51-58, 1990
 A;Title: Plasmodium yoelii nigeriensis circumsporozoite gene structure and its implicati
 A;Reference number: A44969; MUID:91148645; PMID:2290446
 A;Accession: A44969
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-264 <COL>
 A;Cross-references: GB:M32350
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 F;190-242/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 2; Length 264;
 Best Local Similarity 47.4%; Pred. No. 3.9; Indels 0; Gaps 0;
 Matches 9; Conservative 7; Mismatches 3

QY 2 EKXIANEXKASSVFNVNS 20
 Db 232 DTEICKMDKCSIFNIVSN 250

RESULT 8
 OZZQMB
 Circumsporozoite protein precursor - Plasmodium berghei (strain NK65)
 N;Alternate names: sporozoite surface antigen
 C;Species: Plasmodium berghei
 C;Date: 30-Sep-1987 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999
 C;Accession: A44948; A25083; S13446
 R;Janat, D.E.
 Mol. Biochem. Parasitol. 39, 151-154, 1990
 A;Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and NK65
 A;Reference number: A44948; MUID:90158693; PMID:2406593
 A;Accession: A44948
 A;Molecule type: DNA
 A;Residues: 1-332 <LAN>
 A;Cross-references: GB:M28887
 R;Eichinger, D.J.; Arnot, D.E.; Tam, J.P.; Nussenzweig, V.; Enea, V.
 Mol. Cell. Biol. 6, 3965-3972, 1986
 A;Title: Circumsporozoite protein of Plasmodium berghei: gene cloning and identification
 A;Reference number: A25083; MUID:87089740; PMID:2432395
 A;Accession: A25083
 A;Molecule type: DNA
 A;Residues: 1-26, 'I', 28-68, 'PMLRR', 75-126, 'P', 128-134, 'PPPNANDP', 135-332 <EIC>
 A;Cross-references: GB:M14135; NID:gl60245; PIDN:AAA29577.1; PID:gl60246
 R;Weber, J.L.; Egan, J.B.; Lyon, J.A.; Wirtz, R.A.; Charoenvit, Y.; Maloy, W.L.; Hockmeyer
 Exp. Parasitol. 63, 295-300, 1987
 A;Title: Plasmodium berghei: cloning of the circumsporozoite protein gene.
 A;Reference number: S13446; MUID:87218962; PMID:3556207
 A;Accession: S13446
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 61-122, 'A', 124-332 <WEB>

A;Cross-references: GB:M25445; NID:gl60177; PIDN:AAA29531.1; PID:gl60178
 C;Comment: There are three distinct regions in the mature circumsporozoite protein, th
 obic membrane-anchoring sequence.
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C;Keywords: sporozoite; surface antigen; tandem repeat
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-332/Product: circumsporozoite protein #status predicted <MAT>
 F;94-189/Region: 8-residue repeats
 F;199-230/Region: 2-residue repeats
 F;258-310/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 332;
 Best Local Similarity 47.4%; Pred. No. 5; Indels 3; Gaps 0;
 Matches 9; Conservative 7; Mismatches 3

QY 2 EKXIANEXKASSVFNVNS 20
 Db 300 DTEICKMDKCSIFNIVSN 318

RESULT 9
 OZZQMB
 Circumsporozoite protein precursor - Plasmodium berghei (strain ANKA clone 2.341)
 N;Alternate names: sporozoite surface antigen
 C;Species: Plasmodium berghei
 C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
 C;Accession: S07873; S12571
 R;Lockyer, M.J.; Davies, C.S.; Suhrbier, A.; Sinden, R.E.
 Nucleic Acids Res. 18, 376, 1990
 A;Title: Nucleotide sequence of the Plasmodium berghei circumsporozoite protein gene f
 A;Reference number: S07873; MUID:90221834; PMID:2183186
 A;Accession: S07873
 A;Molecule type: DNA
 A;Residues: 1-348 <LOC>
 A;Cross-references: EMBL:X17606
 R;Lockyer, M.J.
 Submitted to the EMBL Data Library, November 1989
 A;Reference number: S12571
 A;Accession: S12571
 A;Molecule type: DNA
 A;Residues: 1-59, 'I', 61-81, 83-348 <LOC2>
 A;Cross-references: EMBL:X17606; NID:g9784; PIDN:CAA35608.1; PID:g9785
 C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
 C;Keywords: tandem repeat
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-348/Product: circumsporozoite protein #status predicted <MAT>
 F;94-205/Region: 8-residue repeats
 F;215-247/Region: 2-residue repeats
 F;274-326/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 348;
 Best Local Similarity 47.4%; Pred. No. 5.3; Indels 3; Gaps 0;
 Matches 9; Conservative 7; Mismatches 3

QY 2 EKXIANEXKASSVFNVNS 20
 Db 316 DTEICKMDKCSIFNIVSN 334

RESULT 10
 AD2528
 Integrase-recombinase protein [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC712
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AD2528
 R;Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriugu
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AD2528
 A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-360 <KUR>
A;Cross-references: GB:AP003602; PIDN:BA877154.1; PID:G17134596; GSPDB:GN00181
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: air7511
A;Genome: plasmid

Query Match 32.5%; Score 55; DB 2; Length 360;
Best Local Similarity 36.0%; Pred. No. 5.5;
Matches 9; Conservative 9; Mismatches 0; Gaps 0;

Qy 2 BKIKAKMEKASSVFNVNNSPSLHW 26
:::|||||::: : : : :
Db 164 EREIAKLEKAAVLDQHN'NMHW 188

RESULT 11
OZZQMY
gonadoliberin precursor - Plasmodium yoelii
N;Alternate names: sporozoite surface antigen
C;Species: Plasmodium yoelii
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: A26271
R;Lab: A.A.; de la Cruz, V.F.; Welsh, J.A.; Charoenvit, Y.; Maloy, W.L.; McCutchan, T.F.
J; Biol. Chem. 262, 2937-2940, 1987
A;Title: Structure of the gene encoding the circumsporozoite protein of Plasmodium yoelii
A;Reference number: A26271; MUID:87137555; PMID:3102479
A;Accession: A26271
A;Molecule type: DNA
A;Residues: 1-367 <LAL>
A;Cross-references: GB:J02695; NID:G160222; PIDN:AAA29558.1; PID:G160223
C;Comment: There are three distinct regions in the mature circumsporozoite protein, the obic membrane-anchoring sequence.
C;Comment: There are 15 copies of a 6-residue repeat and 8 copies of a 4-residue repeat.
C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C;Keywords: sporozoite; surface antigen; tandem repeat
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-367/Product: circumsporozoite protein #status predicted <MAT>
F;139-228/Region: 6-residue repeats (Q-G-P-G-A-P)
F;229-260/Region: 4-residue repeats (Q-Q-P-P)
F;233-345/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 32.5%; Score 55; DB 1; Length 367;
Best Local Similarity 47.4%; Pred. No. 5.6;
Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EKTIKMEKASSVFNVNNS 20
:::|||||::: : : : :
Db 335 DTEICKMDKSSIFNIVSN 353

RESULT 12
178541
gonadoliberin precursor - rhesus macaque (fragment)
N;Alternate names: luteinizing hormone releasing hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C;Accession: I78541
R;Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A;Title: Developmental expression of the genes encoding transforming growth factor alpha
A;Reference number: 158134; MUID:95124501; PMID:7545971
A;Accession: I78541
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-67 <RES>
A;Cross-references: GB:S75918; NID:G912831; PIDN:AA833096.1; PID:G912832
C;Superfamily: gonadoliberin

Query Match 32.0%; Score 54; DB 2; Length 67;
Best Local Similarity 90.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 23 SLHWSYGLRP 32
| | | | | | | | | |
Db 5 SQHWSYGLRP 14

RESULT 13

RHMSG

gonadoliberin precursor - mouse
N;Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N;Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C;Accession: A47578
R;Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolic
Science 234, 1366-1371, 1986
A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible f
A;Reference number: A47578; MUID:87069928; PMID:3024317
A;Accession: A47578
A;Molecule type: DNA
A;Residues: 1-90 <MAS>
A;Cross-references: EMBL:M14872; NID:G193576; PIDN:AAA37717.1; PID:G387175
C;Genetics:
A;Insertions: 45/3; 77/3
C;Function:
A;Description: Gonadoliberin stimulates pituitary secretion of lutropin and follitropin
A;Note: Gonadoliberin-associated protein may have prolactin release inhibiting activity
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1-23/Domain: signal sequence #status predicted <SIG>
F;22-31/Product: gonadoliberin #status predicted <GLB>
F;35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F;22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predi
F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gl

Query Match 32.0%; Score 54; DB 1; Length 90;

Best Local Similarity 90.0%; Pred. No. 1.6; Mismatches 0; Indels 1; Gaps 0;

Qy 23 SLHWSYGLRP 32
| | | | | | | | | |
Db 21 SQHWSYGLRP 30

RESULT 14

RHUG

gonadoliberin precursor [validated] - human
N;Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing
N;Contains: gonadoliberin-associated protein (GAP); progonaoliberin
C;Species: Homo sapiens (man)
C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
A;Accession: S05308; A26173; A93342; A90108; A01410; S45718
R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone g
A;Reference number: S05308; MUID:89366682; PMID:2671939
A;Accession: S05308
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-92 <HAY>
A;Cross-references: EMBL:X15215; NID:G31955; PIDN:CAA33285.1; PID:G31956
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonado
A;Reference number: A94030; MUID:86094338; PMID:2867548
A;Accession: A26173
A;Molecule type: mRNA
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12578; NID:G183418; PIDN:AAA35916.1; PID:G386749
R;Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A;Title: Characterization of cDNA for precursor of human luteinizing hormone releasing

A:Reference number: A93342; MUID:85012739; PMID:6090951
 A:Accession: A93342
 A:Molecule type: mRNA
 A:Residues: 1-15, 'S', 17-92 <SEE>
 A:Cross-references: GB:X01059; NID:G34356; PIDN:CAA25526.1; PID:G34357
 A:Experimental source: placenta
 R:Tan, L.; Rouseau, P.
 Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
 A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
 A:Reference number: A90108; MUID:83126573; PMID:6760865
 A:Accession: A90108
 A:Molecule type: protein
 A:Residues: 24-33 <TAN>
 A:Experimental source: placental trophoblasts
 R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterda
 FEBS Lett. 346, 203-206, 1994
 A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
 A:Reference number: S45718; MUID:94283597; PMID:8013634
 A:Contents: annotation; degradation pathway of synthetic hormone
 C:Genetics:
 A:Gene: GDB:GNRH; LHRH; GRH
 A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
 A:Map position: 8p21-8p11.2
 A:Introns: 47/3; 75/3
 C:Function:
 A:Description: gonadoliberein stimulates pituitary secretion of lutropin and follitropin
 A:Note: gonadoliberein-associated protein may have prolactin release inhibiting activity
 C:Superfamily: gonadoliberein
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
 F:1-23/Domain: signal sequence #status Predicted <SIG>
 F:24-32/Product: progadoliberein #status Predicted <PGN>
 F:37-92/Product: gonadoliberein-associated protein #status Predicted <MAP>
 F:24/Modified site: pyroglutamic acid (Gln) (in mature form) #status experimen
 F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 32.0%; Score 54; DB 1; Length 92;
 Best Local Similarity 90.0%; Pred No. 1.6;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 23 SLHWSYGLRP 32
 | | | | | | | | | |
 DB 23 SQHWSYGLRP 32

RESULT 15
 RHPGG
 Gonadoliberein - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
 C:Accession: A01411
 R:Baba, Y.; Matsuo, H.; Schally, A.V.
 Biochem. Biophys. Res. Commun. 44, 459-463, 1971
 A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the
 A:Reference number: A90172; MUID:72114303; PMID:4946067
 A:Accession: A01411
 A:Molecule type: protein
 A:Residues: 1-10 <BAB>
 R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
 Biochem. Biophys. Res. Commun. 45, 822-827, 1971
 A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase metho
 A:Reference number: A90176; MUID:72065376; PMID:4942726
 A:Contents: annotation; synthesis
 A:Note: the synthetic and natural hormones have the same physicochemical and biological
 R:Baba, Y.; Arimura, A.; Schally, A.V.
 Biochem. Biophys. Res. Commun. 45, 483-487, 1971
 A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
 A:Reference number: A90175; MUID:72117544; PMID:4946275
 A:Contents: annotation
 A:Note: Trp-3 appears to be essential for biological activity
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
 C:Superfamily: gonadoliberein
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyroglutamic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.8%; Score 52; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 HWSYGLRP 32
 | | | | | | | | | |
 DB 2 HWSYGLRP 9

Search completed: March 10, 2004, 09:16:48
 Job time : 11.3312 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 22.0078 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QVIRANSKFIGITELGSLHWSYGLRPX 28

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 147 | 99.3 | 28 | 9 | US-09-848-834A-11 |
| 2 | 136 | 91.9 | 46 | 9 | US-09-848-834A-19 |
| 3 | 112.5 | 76.0 | 27 | 14 | US-10-076-674-7 |
| 4 | 112.5 | 76.0 | 27 | 15 | US-10-355-161A-7 |
| 5 | 82 | 55.4 | 158 | 14 | US-10-297-942-16 |
| 6 | 81.5 | 55.1 | 31 | 14 | US-10-237-656-1 |
| 7 | 81.5 | 55.1 | 33 | 9 | US-09-848-834A-12 |
| 8 | 81 | 54.7 | 194 | 14 | US-10-295-074-47 |
| 9 | 80 | 54.1 | 32 | 14 | US-10-237-656-5 |
| 10 | 80 | 54.1 | 137 | 14 | US-10-237-656-3 |
| 11 | 79 | 53.4 | 194 | 14 | US-10-295-074-46 |
| 12 | 79 | 53.4 | 285 | 14 | US-10-295-074-11 |
| 13 | 78 | 52.7 | 50 | 9 | US-09-943-548-8 |
| 14 | 78 | 52.7 | 50 | 14 | US-10-339-522-8 |
| 15 | 78 | 52.7 | 158 | 14 | US-10-297-942-14 |

| | | | | | |
|----|------|------|-----|----|--------------------|
| 16 | 78 | 52.7 | 285 | 14 | US-10-295-074-9 |
| 17 | 78 | 52.7 | 287 | 14 | US-10-295-074-15 |
| 18 | 77 | 52.0 | 24 | 14 | US-10-128-711-110 |
| 19 | 76.5 | 51.7 | 51 | 9 | US-09-848-834A-20 |
| 20 | 76 | 51.4 | 31 | 9 | US-09-943-548-2 |
| 21 | 76 | 51.4 | 31 | 14 | US-10-339-522-2 |
| 22 | 76 | 51.4 | 158 | 14 | US-10-297-942-4 |
| 23 | 76 | 51.4 | 158 | 14 | US-10-297-942-6 |
| 24 | 74 | 50.0 | 15 | 9 | US-09-862-849-2 |
| 25 | 74 | 50.0 | 15 | 9 | US-09-785-215-4 |
| 26 | 74 | 50.0 | 15 | 10 | US-09-405-986-1 |
| 27 | 74 | 50.0 | 15 | 14 | US-10-204-362-4 |
| 28 | 74 | 50.0 | 15 | 14 | US-10-223-711-7 |
| 29 | 74 | 50.0 | 15 | 14 | US-10-237-656-13 |
| 30 | 74 | 50.0 | 15 | 14 | US-10-223-809A-4 |
| 31 | 74 | 50.0 | 15 | 14 | US-10-261-446-19 |
| 32 | 74 | 50.0 | 15 | 14 | US-10-239-313A-618 |
| 33 | 74 | 50.0 | 15 | 14 | US-10-295-074-3 |
| 34 | 74 | 50.0 | 15 | 15 | US-10-372-111-7 |
| 35 | 74 | 50.0 | 16 | 9 | US-09-848-834A-2 |
| 36 | 74 | 50.0 | 16 | 10 | US-09-930-915A-64 |
| 37 | 74 | 50.0 | 16 | 14 | US-10-082-014-145 |
| 38 | 74 | 50.0 | 16 | 14 | US-10-372-076-175 |
| 39 | 74 | 50.0 | 17 | 10 | US-09-865-294-3 |
| 40 | 74 | 50.0 | 17 | 14 | US-10-239-313A-619 |
| 41 | 74 | 50.0 | 17 | 15 | US-10-346-563-23 |
| 42 | 74 | 50.0 | 17 | 15 | US-10-411-544-5 |
| 43 | 74 | 50.0 | 17 | 15 | US-10-321-717-23 |
| 44 | 74 | 50.0 | 19 | 14 | US-10-239-313A-620 |
| 45 | 74 | 50.0 | 29 | 9 | US-09-732-754-1 |

ALIGNMENTS

RESULT 1

US-09-848-834A-11
; Sequence 11. Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 11
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of t
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to a
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated-glutamine
; NAME/KEY: MOD RES
; LOCATION: (28)..(28)
; OTHER INFORMATION: Amidated-glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(15)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (16)..(19)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (20)..(28)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.

Query Match 55.4%; Score 82; DB 14; Length 158;

```
Best Local Similarity 89.5%; Pred. No. 0.00024;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITELGSPSL 19
Db 115 QYKANSKFIGITELGDRLL 133

RESULT 6
US-10-237-656-1
; Sequence 1, Application US/10237656
; Publication No. US20030152581A1
; GENERAL INFORMATION:
; APPLICANT: SAINT-REMY, Jean-Marie
; APPLICANT: JACQUEMIN, Marc
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF AL
; FILE REFERENCE: 2002-0771/LC/01699
; CURRENT APPLICATION NUMBER: US/10/237,656
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 09/362,731
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 2.1
; SEQ ID NO 1
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
US-10-237-656-1

Query Match 55.1%; Score 81.5; DB 14; Length 31;
Best Local Similarity 61.3%; Pred. No. 5.1e-05;
Matches 19; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

Qy 1 QYKANSKFIGITELG-----PSLHWS 22
Db 1 QYKANSKFIGITELGHHKIKVLPGCHGS 31

RESULT 7
US-09-848-834A-12
; Sequence 12, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of the
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino ac
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)_(1)
; OTHER INFORMATION: Amidated aspartic acid
; NAME/KEY: MOD RES
; LOCATION: (33)_(33)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)_(20)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
; OTHER INFORMATION: (CSP) protein
```

```
; NAME/KEY: PEPTIDE
; LOCATION: (21)_(24)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (25)_(33)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-12

Query Match 55.1%; Score 81.5; DB 9; Length 33;
Best Local Similarity 66.7%; Pred. No. 5.5e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Qy 4 KANSKFIGITELGPSLHWSYGURP 27
Db 10 KASSVF-NVNSGPSLHWSYGURP 32

RESULT 8
US-10-295-074-47
; Sequence 47, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 47
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
; NAME/KEY: MUTAGEN
; LOCATION: (110)_(130)
; OTHER INFORMATION: Tetanus toxoid P30 epitope
; FEATURE:
; NAME/KEY: MUTAGEN
; LOCATION: (131)_(145)
; OTHER INFORMATION: Tetanus toxoid P2 epitope
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (2)_(109)
; OTHER INFORMATION: hTNF amino acids 1-108
; NAME/KEY: MISC FEATURE
; LOCATION: (146)_(194)
; OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-47

Query Match 54.7%; Score 81; DB 14; Length 194;
Best Local Similarity 76.2%; Pred. No. 0.00042;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITELGPSLHW 21
Db 131 QYKANSKFIGITELAEAXPW 151

RESULT 9
US-10-237-656-5
; Sequence 5, Application US/10237656
; Publication No. US20030152581A1
; GENERAL INFORMATION:
; APPLICANT: SAINT-REMY, Jean-Marie
; APPLICANT: JACQUEMIN, Marc
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF
; FILE REFERENCE: 2002-0771/LC/01699
; CURRENT APPLICATION NUMBER: US/10/237,656
; CURRENT FILING DATE: 2002-09-10
```



```
; PRIOR APPLICATION NUMBER: 09/362,731
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Version 2.1
; SEQ ID NO 5
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
US-10-237-656-5

Query Match          54.1%; Score 80; DB 14; Length 32;
Best Local Similarity 100.0%; Pred. No. 8.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
DB 1 QYIKANSKFIGITELG 16

RESULT 10
US-10-237-656-3
; Sequence 3, Application US/10237656
; Publication No. US20030152581A1
; GENERAL INFORMATION:
; APPLICANT: SAINT-REMY, Jean-Marie
; APPLICANT: JACQUEMIN, Marc
; TITLE OF INVENTION: COMPOUND AND METHOD FOR THE PREVENTION AND/OR THE TREATMENT OF AL
; FILE REFERENCE: 2002-0771/LC/01699
; CURRENT APPLICATION NUMBER: US/10/237,656
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 09/362,731
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Version 2.1
; SEQ ID NO 3
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide
US-10-237-656-3

Query Match          54.1%; Score 80; DB 14; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELG 16
DB 2 QYIKANSKFIGITELG 17

RESULT 11
US-10-295-074-46
; Sequence 46, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 46
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
; NAME/KEY: MUTAGEN
```

```
; LOCATION: (110)..(124)
; OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 2)
; FEATURE:
; NAME/KEY: MUTAGEN
; LOCATION: (125)..(145)
; OTHER INFORMATION: Tetanus toxoid P30 epitope (SEQ ID NO: 3)
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (145)..(194)
; OTHER INFORMATION: hTNF amino acids 1-108
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (145)..(194)
; OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-46

Query Match          53.4%; Score 79; DB 14; Length 194;
Best Local Similarity 69.2%; Pred. No. 0.00084;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGPSLHWSYGLR 26
DB 110 QYIKANSKFIGITELFNNFTVFWLR 135

RESULT 12
US-10-295-074-11
; Sequence 11, Application US/10295074
; Publication No. US20030185845A1
; GENERAL INFORMATION:
; APPLICANT: Pharmexa A/S
; TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
; FILE REFERENCE: P1013DK00
; CURRENT APPLICATION NUMBER: US/10/295,074
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 11
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: 2 human IL5 monomers joined by P2 and P30 epitopes
US-10-295-074-11

Query Match          53.4%; Score 79; DB 14; Length 285;
Best Local Similarity 69.2%; Pred. No. 0.0013;
Matches 18; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGPSLHWSYGLR 26
DB 135 QYIKANSKFIGITELFNNFTVFWLR 160

RESULT 13
US-09-943-548-8
; Sequence 8, Application US/09943548
; Patent No. US20020042364A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-1; TCS-411.1P US-2
; CURRENT APPLICATION NUMBER: US/09/943,548
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.1
```

```
; SEQ ID NO 8
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
US-09-943-548-8

Query Match          52.7%; Score 78; DB 9; Length 50;
Best Local Similarity 94.1%; Pred. No. 0.00028;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGP 17
   |||||
Db 2 QYIKANSKFIGITELFP 18
   |||||

RESULT 14
US-10-339-522-8
; Sequence 8, Application US/10339522
; Publication No. US20030108559A1
; GENERAL INFORMATION:
; APPLICANT: Rittershaus, Charles W.
; APPLICANT: Thomas, Lawrence J.
; TITLE OF INVENTION: MODULATION OF CHOLESTERYL ESTER TRANSFER PROTEIN (CETP) ACTIVITY
; FILE REFERENCE: TCS-411.1P US-3
; CURRENT APPLICATION NUMBER: US/10/339,522
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 08/432,483
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: PCT/US96/06147
; PRIOR FILING DATE: 1996-05-01
; PRIOR APPLICATION NUMBER: 08/945,289
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 09/943,334
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/943,548
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: vaccine peptide of the invention
US-10-339-522-8

Query Match          52.7%; Score 78; DB 14; Length 50;
Best Local Similarity 94.1%; Pred. No. 0.00028;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITELGP 17
   |||||
Db 2 QYIKANSKFIGITELFP 18
   |||||

RESULT 15
US-10-297-942-14
; Sequence 14, Application US/10297942
; Publication No. US20030185816A1
; GENERAL INFORMATION:
; APPLICANT: Ferring BV
; TITLE OF INVENTION: Solubilised Protein Vaccines
; FILE REFERENCE: P68445US0
; CURRENT APPLICATION NUMBER: US/10/297,942
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: PCT/DK01/00431
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: DK PA 2000 00966
; PRIOR FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 14
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-297-942-14

Query Match          52.7%; Score 78; DB 14; Length 158;
Best Local Similarity 81.0%; Pred. No. 0.00095;
Matches 17; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 QYIKANSKFIGITELGPSLHW 21
   |||||
Db 11 QYIKANSKFIGITEL--QLQW 29
   |||||

Search completed: March 10, 2004, 10:25:48
Job time : 22.0078 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 42.1634 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-11
Perfect score: 148
Sequence: 1 QYIKANKFQIGITELGRLHSYGLRPX 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000a:*
- 4: Geneseq2001s:*
- 5: Geneseq2002a:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|------------|--------------------|
| 1 | 147 | 99.3 | 28 | 5 AAU11422 | AAU11422 Synthetic |
| 2 | 136 | 91.9 | 46 | 5 AAU11430 | AAU11430 Synthetic |
| 3 | 119.5 | 80.7 | 47 | 2 AAR62723 | AAR62723 LHRH-cont |
| 4 | 112.5 | 76.0 | 27 | 2 AAR62701 | AAR62701 LHRH-cont |
| 5 | 112.5 | 76.0 | 27 | 7 ADD89947 | ADD89947 LHRH-pept |
| 6 | 86 | 58.1 | 750 | 3 AAY92639 | AAY92639 Mutant hu |
| 7 | 82 | 55.4 | 158 | 2 AAW81329 | AAW81329 TNF2-4, a |
| 8 | 82 | 55.4 | 158 | 5 ABB07281 | ABB07281 Human TNF |
| 9 | 82 | 55.4 | 693 | 3 AAY92649 | AAY92649 Mutant hu |
| 10 | 82 | 55.4 | 750 | 3 AAY92630 | AAY92630 Mutant hu |
| 11 | 82 | 55.4 | 750 | 3 AAY92646 | AAY92646 Mutant hu |
| 12 | 82 | 55.4 | 750 | 3 AAY92641 | AAY92641 Mutant hu |
| 13 | 81.5 | 55.1 | 31 | 3 AAY82632 | AAY82632 Tetanus t |
| 14 | 81.5 | 55.1 | 33 | 5 AAU11423 | AAU11423 Synthetic |
| 15 | 81.5 | 54.7 | 194 | 6 AAO30489 | AAO30489 Human TNF |
| 16 | 80.5 | 54.4 | 708 | 7 ABR82479 | ABR82479 Modified |
| 17 | 80 | 54.1 | 29 | 2 AAR83561 | AAR83561 IGE CH4 r |
| 18 | 80 | 54.1 | 32 | 3 AAY82636 | AAY82636 Tetanus t |
| 19 | 80 | 54.1 | 37 | 2 AAR65389 | AAR65389 Universal |
| 20 | 80 | 54.1 | 37 | 2 AAR65383 | AAR65383 Universal |
| 21 | 80 | 54.1 | 109 | 4 AAB20147 | AAB20147 Growth di |
| 22 | 80 | 54.1 | 137 | 3 AAY82634 | AAY82634 Tetanus t |
| 23 | 79 | 53.4 | 43 | 4 AAB49076 | AAB49076 Amyloid b |
| 24 | 79 | 53.4 | 43 | 4 AAB46177 | AAB46177 Tetanus t |
| 25 | 79 | 53.4 | 72 | 4 AAB46190 | AAB46190 Tetanus t |

| | | | | | |
|----|------|------|-----|-------------|---------------------|
| 26 | 79 | 53.4 | 136 | 4 AAB49089 | AAB49089 Amyloid b |
| 27 | 79 | 53.4 | 145 | 3 AAB45530 | AAB45530 Modified |
| 28 | 79 | 53.4 | 147 | 3 AAB45522 | AAB45522 Modified |
| 29 | 79 | 53.4 | 194 | 6 AAO30488 | AAO30488 Human TNF |
| 30 | 79 | 53.4 | 254 | 4 AAB20152 | AAB20152 Growth di |
| 31 | 79 | 53.4 | 285 | 6 AAO30458 | AAO30458 hIL5-P2-P |
| 32 | 79 | 53.4 | 717 | 7 ABR82478 | ABR82478 Modified |
| 33 | 79 | 53.4 | 750 | 3 AAY92627 | AAY92627 Mutant hu |
| 34 | 78 | 52.7 | 30 | 2 AAR44398 | AAR44398 HIV antiq |
| 35 | 78 | 52.7 | 50 | 2 AAW06131 | AAW06131 Anti-choi |
| 36 | 78 | 52.7 | 158 | 2 AAB01327 | AAB01327 TNF2-1, a |
| 37 | 78 | 52.7 | 158 | 5 ABB07280 | ABB07280 Human TNF |
| 38 | 78 | 52.7 | 182 | 3 AAY84424 | AAY84424 An osteop |
| 39 | 78 | 52.7 | 285 | 6 AAO30457 | AAO30457 hIL5-P30- |
| 40 | 78 | 52.7 | 287 | 6 AAO30460 | AAO30460 hIL5-37 v |
| 41 | 78 | 52.7 | 537 | 7 ABR82481 | ABR82481 Truncated |
| 42 | 78 | 52.7 | 713 | 7 ABR82480 | ABR82480 Modified |
| 43 | 77 | 52.0 | 24 | 5 ABR251509 | ABR251509 HBV antiq |
| 44 | 77 | 52.0 | 27 | 5 ABR251516 | ABR251516 HBV antiq |
| 45 | 76.5 | 51.7 | 51 | 5 AAU11431 | AAU11431 Synthetic |

ALIGNMENTS

RESULT 1

AAU11422
ID AAU11422 standard; peptide; 28 AA.

XX AC AAU11422;

XX DT 12-MAR-2002 (first entry)

XX DE Synthetic immunogen peptide 3.

XX KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;

XX KW luteinising hormone releasing hormone; LHRH; contraceptive;

XX KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

XX KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

XX KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX OS Clostridium tetani.

OS Mammalia.

OS Synthetic.

OS OS Chimeric.

XX FH Key Location/Qualifiers

FT Peptide 1..15

FT Peptide /note= "Tetanus toxoid sequence (830-844 aa)"

FT Peptide 16..19

FT Peptide /note= "Spacer peptide"

FT Peptide 20..28

FT Modified-site 28 /note= "Gonadotrophin releasing hormone epitope"

XX /note= "Amidated glycine or glycineamide"

PN WO200185763-A2.

XX 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US014363.

XX 05-MAY-2000; 2000US-020328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against

XX Gonadotropin releasing hormone; comprises fusion peptide having

XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope

PT or its analog.
 XX
 PS Claim 11; Page 8; 43pp; English.
 XX
 CC The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 SQ Sequence 28 AA;
 Query Match 99.3%; Score 147; DB 5; Length 28;
 Best Local Similarity 100.0%; Pred. No. 2.1e-15;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITELGPSLHWSYGLRP 27
 DB 1 QYIKANSKFIGITELGPSLHWSYGLRP 27
 RESULT 2
 AAU11430
 ID AAU11430 standard; peptide; 46 AA.
 AC AAU11430;
 XX
 DT 12-MAR-2002 (first entry)
 DE Synthetic immunogen peptide 11.
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 OS Clostridium tetani.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..10
 FT /note= "Gonadotropin releasing hormone epitope (1..10
 FT aa)"
 FT Misc-difference 1
 FT /label= OTHER
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
 FT Peptide 11..16
 FT /note= "Spacer peptide"
 FT Peptide 17..31
 FT /note= "Tetanus toxoid (830-844 aa)"
 FT Peptide 32..37
 FT /note= "Spacer peptide"
 FT Peptide 38..46
 FT /note= "Gonadotropin releasing hormone epitope (2-10
 FT aa)"
 FT Modified-site 46
 FT /note= "Amidated glycine or glycineamide"
 XX
 PN WO200185763-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 04-MAY-2001; 2001WO-US014363.

XX 05-MAY-2000; 2000US-0202328P.
 PR (AFHT-) APHTON CORP.
 PA
 XX Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 XX gonadotropin releasing hormone, comprises fusion peptide having
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 XX or its analog.
 XX Claim 11; Page 12; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 SQ Sequence 46 AA;
 Query Match 91.9%; Score 136; DB 5; Length 46;
 Best Local Similarity 93.1%; Pred. No. 1.8e-13;
 Matches 27; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 QY 1 QYIKANSKFIGITEL--GPSLHWSYGLRP 27
 DB 17 QYIKANSKFIGITELSSGPSLHWSYGLRP 45
 RESULT 3
 AAR62723
 ID AAR62723 standard; peptide; 47 AA.
 XX AAR62723;
 XX 25-MAR-2003 (revised)
 DT 17-SEP-1995 (first entry)
 XX LHRH-containing immunogenic peptide.
 XX Helper T cell epitope; universal immune stimulator; invasive; haptens;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate; toxin.
 KW androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Domain 1..16 "invasin domain"
 FT /note= "invasin domain"
 FT Domain 19..35
 FT /note= "tetanus toxin helper T cell epitope"
 FT Domain 38..47
 FT /note= "LHRH haptens"
 XX
 PN WO9425060-A1.
 XX
 PD 10-NOV-1994.
 XX
 PF 28-APR-1994; 94WO-US004832.
 XX
 PD 27-APR-1993; 93US-00057166.
 PR 14-APR-1994; 94US-00229275.

```

XX (LADD/) LADD A E.
PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX
PI Ladd AE, Wang CY, Zamb T;
XX
DR WPI; 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT suppress LHRH activity in males and females.
XX
PS Claim 8; Page 88; 213pp; English.
XX
CC Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The stimulator
CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC immune response to the coupled peptide in members of a heterogeneous
CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC sequence from the invasive protein of Yersinia. Spacer amino acid
CC sequences (e.g. Gly-Gly) can be provided between the invasive and Th
CC domains and between the immune stimulator and hapten components. When the
CC hapten is LHRH, then optionally the invasive domain can be omitted from
CC the immune stimulator component. The present sequence represents an LHRH-
CC containing immunogenic peptide as above which can be used as a potent
CC vaccine for treating e.g. prostatic hyperplasia, androgen-dependent
CC carcinoma, prostatic carcinoma, testicular carcinoma, endometriosis,
CC benign uterine tumours, recurrent functional ovarian cysts, (severe)
CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
CC induction of infertility. (Updated on 25-MAR-2003 to correct FN field.)
XX
SQ Sequence 47 AA;
Query Match 80.7%; Score 119.5; DB 2; Length 47;
Best Local Similarity 88.9%; Pred. No. 6.7e-11;
Matches 24; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 QYIKANSKFIGITELGSLHWSYGLRP 27
DB 21 QYIKANSKFIGITELGGE-HWSYGLRP 46

RESULT 4
AAR62701
ID AAR62701 standard; peptide; 27 AA.
XX
AC AAR62701;
XX
XX 25-MAR-2003 (revised)
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
XX Helper T cell epitope; universal immune stimulator; invasive; hapten;
XX vaccine; LHRH; luteinising hormone releasing hormone; prostate;
XX androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Domain 1..17
FT Domain /note="tetanus toxin helper T cell epitope"
FT Domain 18..27
FT Domain /note="LHRH hapten"
XX
XX WO9425060-A1.
XX
XX 10-NOV-1994.
XX
XX 28-APR-1994; 94WO-US004832.
XX

```

```

PR 27-APR-1993; 93US-00057166.
PR 14-APR-1994; 94US-00229275.
XX
PA (LADD/) LADD A E.
PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX
PI Ladd AE, Wang CY, Zamb T;
XX
DR WPI; 1994-357910/44.
XX
XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT suppress LHRH activity in males and females.
XX
XX Claim 8, 12; Page 84; 213pp; English.
XX
CC Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The stimulator
CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC immune response to the coupled peptide in members of a heterogeneous
CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC sequence from the invasive protein of Yersinia. Spacer amino acid
CC sequences (e.g. Gly-Gly) can be provided between the invasive and Th
CC domains and between the immune stimulator and hapten components. When the
CC hapten is LHRH, then optionally the invasive domain can be omitted from
CC the immune stimulator component. The present sequence represents an LHRH-
CC containing, invasive-free immunogenic peptide as above which can be used
CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
CC endometriosis, benign uterine tumours, recurrent functional ovarian
CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
CC cancer, or for induction of infertility. This sequence is particularly
CC preferred. (Updated on 25-MAR-2003 to correct FN field.)
XX
SQ Sequence 27 AA;
Query Match 76.0%; Score 112.5; DB 2; Length 27;
Best Local Similarity 85.2%; Pred. No. 4.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 QYIKANSKFIGITELGSLHWSYGLRP 27
DB 3 QYIKANSKFIGITELB--HWSYGLRP 26

RESULT 5
ADD89947
ID ADD89947 standard; protein; 27 AA.
XX
AC ADD89947;
XX
XX 29-JAN-2004 (first entry)
DT
XX
DE LHRH peptide used in immunostimulant complex for prostate cancer vaccine.
XX
XX Immunostimulant; vaccine; human; immunogen; LHRH; immunotherapy;
XX prostate cancer.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX WO2003068169-A2.
XX
XX 21-AUG-2003.
XX
XX 14-FEB-2003; 2003WO-US004711.
XX
XX 14-FEB-2002; 2002US-00076674.
XX
XX 31-JAN-2003; 2003US-00076674.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
PA

```

XX PI Sokoll KK;
XX WPI; 2003-778890/73.
XX Stabilized immunostimulating complex, useful for vaccination, e.g.
XX against human immune deficiency viruses, comprises cationic peptide
XX immunogen and anionic oligonucleotide.
XX
XX Claim 17; SEQ ID NO 7; 159pp; English.
XX
XX The present sequence is that of a synthetic immunogenic peptide derived
XX from human LHRH. This is an example of peptides that can be used in
XX claimed immunostimulatory complexes of the invention that are
XX specifically adapted to act as adjuvant and as peptide immunogen
XX stabiliser. The complexes comprise a CpG oligonucleotide and a
XX biologically active peptide immunogen. The complex is particulate and can
XX efficiently present peptide immunogens to the cells of the immune system
XX to produce an immune response. The complexes may be prepared with various
XX ratios of peptides to CpG oligonucleotides to provide different physical
XX properties, such as the size of the microparticle. An immunostimulatory
XX complex comprising the present LHRH derived peptide can be used in a
XX vaccine for prostate cancer.
XX
XX Sequence 27 AA;
SQ
Query Match 76.0%; Score 112.5; DB 7; Length 27;
Best Local Similarity 85.2%; Pred. No. 4.4e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 QYIKANSKFIGITELGSLHWSYGLRP 27
DB 3 QYIKANSKFIGITELGSLHWSYGLRP 26
RESULT 6
ID AAY92639 standard; protein; 750 AA.
XX
XX AAY92639;
DT 10-AUG-2000 (first entry)
XX
XX Mutant human prostate specific membrane antigen construct, hPSM5.1.
XX
XX Prostate specific membrane antigen; immunogenized construct; mutant;
XX vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
XX prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 21..41
XX /label= P30
XX /note= "foreign epitope"
XX Peptide 305..319
XX /label= P2
XX /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
XX 13-APR-2000.
XX
XX 05-OCT-1999; 99WO-DK000525.
XX
XX 05-OCT-1998; 98DK-00001261.
XX 20-OCT-1998; 98US-0105011P.
XX
XX (NEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX Gautam A, Birk P, Karlsson G;
XX

XX DR WPI; 2000-349917/30.
XX Inducing immune responses to weakly immunogenic, tumor associated peptide
XX antigens for the treatment of breast and prostate cancer.
XX
XX Example 1; Page; 220pp; English.
XX
XX AAY92627-49 are mutant immunogenized human prostate specific membrane
XX antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
XX The immunogenic analogues of PSM can be used in the claimed method as an
XX autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX binding regions and cysteine residues involved in disulfide bonds are
XX preserved in the immunogenized forms. The method is used for inducing
XX immune responses against weakly immunogenic cell-associated peptide
XX antigens (PA) such as those associated with cancers (self-proteins), e.g.
XX human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX simultaneous presentation by antigen producing cells (APCs) of the
XX animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
XX group derived from the PA and/or at least 1 B-cell group derived from the
XX cell-associated PA; and (2) at least 1 first T helper cell group which is
XX foreign to the animal. Analogues of human PSM, human Her2 and
XX human/murine FGF8b comprising a substantial part of all known and
XX predicted CTL and B-cell epitopes of the respective PA and including at
XX least one foreign T helper epitope are also claimed. The method is used
XX to treat prostate, prostate/breast or breast cancer when the PA is human
XX PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
XX from the wild type human PSM (AAY92619), which appears on pages 184-187
XX of the specification
SQ
Sequence 750 AA;
Query Match 59.1%; Score 86; DB 3; Length 750;
Best Local Similarity 70.4%; Pred. No. 0.00023;
Matches 19; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 QYIKANSKFIGITELGSLHWSYGLRP 27
DB 305 QYIKANSKFIGITELGSLHWSYGLRP 331
RESULT 7
AAW81329
ID AAW81329 standard; protein; 158 AA.
XX
XX AAW81329;
DT 21-APR-1999 (first entry)
XX
XX TNF2-4, a TNF-alpha analogue.
XX
XX Human tumour necrosis factor-alpha; TNF-alpha; TNF-alpha analogue;
XX vaccine; rheumatoid arthritis; Crohn's disease; ulcerative colitis;
XX cancer; disseminated sclerosis; diabetes; psoriasis; osteoporosis;
XX asthma.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9846642-A1.
XX
XX 22-OCT-1998.
XX
XX 15-APR-1998; 98WO-DK000157.
XX
XX 15-APR-1997; 97DK-00000418.
XX 24-APR-1997; 97US-0044187P.
XX
XX (FERR) FARM LAB FERRING AS.
XX
XX Jensen MR, Mouritsen S, Elsner H, Dalum I;
XX

DR WPI; 1998-594561/50.
 DR N-PSDB; AAV68418.
 XX Modified human tumour necrosis factor-alpha - comprises immunodominant T
 PT cell epitope, useful in vaccines to treat or prevent TNF-associated
 PT diseases, e.g. cancer.
 XX
 XX
 XX Example 1; Page 69-70; 134pp; English.
 PS
 CC The present sequence represents a modified human tumour necrosis factor-
 CC alpha (TNF-alpha) analogue. The analogues have no residual TNF activity
 CC and are immunogenic in a large proportion of the human population (by
 CC using promiscuous epitopes). The TNF-alpha analogue is able to generate,
 CC in humans, neutralizing antibodies to wild-type human TNF alpha, has at
 CC least one fragment of TNF substituted by a peptide containing an
 CC immunodominant T-cell epitope, and at least one TNF-alpha B-cell epitope.
 CC The substitution causes a significant change in the amino acid sequence
 CC of any one of the strands in the front beta-sheet, any of the connecting
 CC loops or any of the B', I or D strands in the back beta-sheet. The TNF-
 CC alpha analogues are used as vaccines for treatment or prevention of
 CC diseases associated with excessive release or activity of TNF-alpha, e.g.
 CC rheumatoid arthritis, Crohn's disease, ulcerative colitis, cancer of any
 CC sort, disseminated sclerosis, diabetes, psoriasis, osteoporosis and
 CC asthma
 XX
 XX Sequence 158 AA;
 SQ
 Query Match 55.4%; Score 82; DB 2; Length 158;
 Best Local Similarity 89.5%; Pred. No. 0.00017;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITELGDSL 19
 DB 115 QYIKANSKFIGITELGDSL 133
 RESULT 8
 ID ABB07281 standard; protein; 158 AA.
 XX
 XX ABB07281;
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human TNF-alpha analogue TNF2-4.
 XX
 KW TNF-alpha; pharmaceutical; vaccine; self-protein; tumour necrosis factor;
 KW cetylpyridinium chloride; immunisation; antiinflammatory; antirheumatic;
 KW antiarthritic; antiulcer; cytostatic; antidiabetic; antipsoriatic;
 KW antiasthmatic; immunomodulator; neuroprotective; osteoprotective; human;
 KW TNF2-4.
 XX
 OS Homo sapiens.
 XX
 XX WO200197837-A1.
 FN
 XX
 PD 27-DEC-2001.
 XX
 XX 20-JUN-2001; 2001WO-DK000431.
 PF
 XX
 PR 21-JUN-2000; 2000DK-00000965.
 PR
 XX
 XX (FERR) FERRING BV.
 PA
 XX
 PI Olesen OF, Balchen T, Bouman MHEM;
 XX
 DR WPI; 2002-114542/15.
 DR N-PSDB; ABA94391.
 DR
 XX Novel vaccine composition for prevention/treatment of self-protein-
 PT mediated pathology such as cancer, diabetes and asthma, comprises
 PT modified immunogenic self-protein and surfactant capable of acting as
 PT solubilizer.

XX Claim 21; Page 46-47; 55pp; English.
 PS
 XX The invention provides a pharmaceutical vaccine composition (I) for the
 CC prevention or treatment of a self-protein-mediated pathology. The
 CC composition comprises at least one modified immunogenic self-protein
 CC (selected from modified TNF-alpha proteins) and a surfactant capable of
 CC acting as a solubilizer. (I) is useful for preventing or treating a self
 CC -protein-mediated pathology such as an inflammatory disease, rheumatoid
 CC arthritis, an inflammatory bowel disease (ulcerative colitis or Crohn's
 CC disease), cancer, cachexia, multiple sclerosis, diabetes, psoriasis,
 CC osteoporosis or asthma. (I) is useful for inducing autoantibodies to a
 CC self-protein such as TNF (tumour necrosis factor)-alpha in a human
 CC subject. (I) comprising cetylpyridinium chloride as a component is useful
 CC for immunisation of a human subject or for treatment of a human
 CC inflammatory disease. The present sequence represents a human TNF-alpha
 CC analogue TNF2-4
 XX
 XX Sequence 158 AA;
 SQ
 Query Match 55.4%; Score 82; DB 5; Length 158;
 Best Local Similarity 89.5%; Pred. No. 0.00017;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 QYIKANSKFIGITELGDSL 19
 DB 115 QYIKANSKFIGITELGDSL 133
 RESULT 9
 ID AAY92649 standard; protein; 693 AA.
 XX
 XX AAY92649;
 AC
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Mutant human PSM antigen splice variant construct, hPSM'10.3.
 XX
 KW Prostate specific membrane antigen; immunogenized construct; mutant;
 KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer; PSM;
 KW prostate cancer; cell-associated peptide antigen; foreign epitope.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT 153..173
 FT Peptide /label= P30
 FT /note= "foreign epitope"
 FT 617..631
 FT Peptide /label= P2
 FT /note= "foreign epitope"
 FT
 XX WO200020027-A2.
 PN
 XX
 XX 13-APR-2000.
 PD
 XX
 XX 05-OCT-1999; 99WO-DK000525.
 PF
 XX
 XX 05-OCT-1998; 98DK-00001261.
 PR
 XX 20-OCT-1998; 98US-0105011P.
 PR
 XX
 XX (MEBI-) M & E BIOTECH AS.
 PA
 XX Steinaa I, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX
 XX WPI; 2000-349917/30.
 DR
 XX Inducing immune responses to weakly immunogenic, tumor associated peptide
 PT antigens for the treatment of breast and prostate cancer.
 PT
 XX

```

PS Example 1; Page: 220pp; English.
XX
CC AAY92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC The immunogenic analogues of PSM can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC binding regions and cysteine residues involved in disulfide bonds are
CC preserved in the immunogenized forms. The method is used for inducing
CC immune responses against weakly immunogenic cell-associated peptide
CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animal's immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
CC from the wild type human PSM (AAY92619), which appears on pages 184-187
CC of the specification
XX
SQ Sequence 693 AA;
Query Match 55.4%; Score 82; DB 3; Length 693;
Best Local Similarity 70.4%; Pred. No. 0.00087;
Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 1 QYIKANSKFGITEL-----GPSLHWSY 23
DB 617 QYIKANSKFGITELHVIYAPSSHNY 643

RESULT 10
AAY92630
ID AAY92630 standard; protein; 750 AA.
XX
AC AAY92630;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant human prostate specific membrane antigen construct, hPSM10.1.
XX
KW Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 21..41
FT /label= P30
FT /note= "foreign epitope"
FT Peptide 674..688
FT /label= P2
FT /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
PN 210..230
XX /label= P30
XX /note= "foreign epitope"
PD FT 674..688
XX Peptide /label= P2
XX /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
PN 13-APR-2000.
XX
PD 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
XX 20-OCT-1998; 98US-0105011P.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

```

```

PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated peptide
XX antigens for the treatment of breast and prostate cancer.
XX
XX Example 1; Page: 220pp; English.
XX
CC AAY92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC The immunogenic analogues of PSM can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC binding regions and cysteine residues involved in disulfide bonds are
CC preserved in the immunogenized forms. The method is used for inducing
CC immune responses against weakly immunogenic cell-associated peptide
CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animal's immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
CC from the wild type human PSM (AAY92619), which appears on pages 184-187
CC of the specification
XX
SQ Sequence 750 AA;
Query Match 55.4%; Score 82; DB 3; Length 750;
Best Local Similarity 70.4%; Pred. No. 0.00095;
Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 1 QYIKANSKFGITEL-----GPSLHWSY 23
DB 674 QYIKANSKFGITELHVIYAPSSHNY 700

RESULT 11
AAY92646
ID AAY92646 standard; protein; 750 AA.
XX
AC AAY92646;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant human prostate specific membrane antigen construct, hPSM10.3.
XX
KW Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 210..230
FT /label= P30
FT /note= "foreign epitope"
FT Peptide 674..688
FT /label= P2
FT /note= "foreign epitope"
XX
XX WO200020027-A2.
XX
PN 13-APR-2000.
XX
PD 05-OCT-1999; 99WO-DK000525.
XX

```



```

XX PR 05-OCT-1998; 98DK-00001261.
XX PR 20-OCT-1998; 98US-0105011P.
XX PA (MEBI-) M & E BIOTECH AS.
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX PR WPI; 2000-349917/30.
XX XX
XX XX Inducing immune responses to weakly immunogenic, tumor associated peptide
XX PT antigens for the treatment of breast and prostate cancer.
XX PS
XX PS Example 1; Page; 220pp; English.
XX CC AAY92627-49 are mutant immunogenized human prostate specific membrane
XX CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
XX CC The immunogenic analogues of PSM can be used in the claimed method as an
XX CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX CC binding regions and cysteine residues involved in disulfide bonds are
XX CC preserved in the immunogenized forms. The method is used for inducing
XX CC immune responses against weakly immunogenic cell-associated peptide
XX CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
XX CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX CC simultaneous presentation by antigen producing cells (APCs) of the
XX CC group derived from the PA and/or at least 1 CTL (cytotoxic T-lymphocyte)
XX CC cell-associated PA; and (2) at least 1 B-cell group derived from the
XX CC foreign to the animal. Analogues of human PSM, human Her2 and
XX CC human/murine FGF8b comprising a substantial part of all known and
XX CC predicted CTL and B-cell epitopes of the respective PA and including at
XX CC least one foreign T helper epitope are also claimed. The method is used
XX CC to treat prostate, prostate/breast or breast cancer when the PA is human
XX CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
XX CC from the wild type human PSM (AAY92619), which appears on pages 184-187
XX CC of the specification
XX SQ Sequence 750 AA;
XX
XX Query Match 55.4%; Score 82; DB 3; Length 750;
XX Best Local Similarity 70.4%; Pred. No. 0.00095;
XX Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
XX
XX Qy 1 QYIKANSKFIGITEL---GPSLHWSY 23
XX Db 674 QYIKANSKFIGITELHVIYAPSHNKY 700
XX
XX RESULT 12
XX AAY92641
XX ID AAY92641 standard; protein; 750 AA.
XX XX
XX AC AAY92641;
XX XX
XX DT 10-AUG-2000 (first entry)
XX DE Mutant human prostate specific membrane antigen construct, hPSM10.0.
XX XX
XX KW Prostate specific membrane antigen; immunogenized construct; mutant;
XX KW vaccination; Cytotoxic T-lymphocyte immunity; breast cancer;
XX KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX XX
XX OS Homo sapiens.
XX OS Synthetic.
XX PH Key Location/Qualifiers
XX FT Peptide 674..688
XX FT /label= P2
XX FT /note= "foreign epitope"
XX XX
XX WO200020027-A2.

```

```

XX PR 13-APR-2000.
XX XX
XX PF 05-OCT-1999; 99WO-DK000525.
XX XX
XX PR 05-OCT-1998; 98DK-00001261.
XX PR 20-OCT-1998; 98US-0105011P.
XX XX
XX PA (MEBI-) M & E BIOTECH AS.
XX XX
XX PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
XX PI Gautam A, Birk P, Karlsson G;
XX XX
XX PR WPI; 2000-349917/30.
XX XX
XX XX Inducing immune responses to weakly immunogenic, tumor associated peptide
XX PT antigens for the treatment of breast and prostate cancer.
XX PS
XX PS Example 1; Page; 220pp; English.
XX CC AAY92627-49 are mutant immunogenized human prostate specific membrane
XX CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
XX CC The immunogenic analogues of PSM can be used in the claimed method as an
XX CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
XX CC binding regions and cysteine residues involved in disulfide bonds are
XX CC preserved in the immunogenized forms. The method is used for inducing
XX CC immune responses against weakly immunogenic cell-associated peptide
XX CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
XX CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
XX CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
XX CC simultaneous presentation by antigen producing cells (APCs) of the
XX CC group derived from the PA and/or at least 1 CTL (cytotoxic T-lymphocyte)
XX CC cell-associated PA; and (2) at least 1 B-cell group derived from the
XX CC foreign to the animal. Analogues of human PSM, human Her2 and
XX CC human/murine FGF8b comprising a substantial part of all known and
XX CC predicted CTL and B-cell epitopes of the respective PA and including at
XX CC least one foreign T helper epitope are also claimed. The method is used
XX CC to treat prostate, prostate/breast or breast cancer when the PA is human
XX CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed
XX CC from the wild type human PSM (AAY92619), which appears on pages 184-187
XX CC of the specification
XX SQ Sequence 750 AA;
XX
XX Query Match 55.4%; Score 82; DB 3; Length 750;
XX Best Local Similarity 70.4%; Pred. No. 0.00095;
XX Matches 19; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
XX
XX Qy 1 QYIKANSKFIGITEL---GPSLHWSY 23
XX Db 674 QYIKANSKFIGITELHVIYAPSHNKY 700
XX
XX RESULT 13
XX AAY92632
XX ID AAY92632 standard; peptide; 31 AA.
XX XX
XX AC AAY92632;
XX XX
XX DT 07-AUG-2000 (first entry)
XX DE Tetanus toxoid T cell epitope and Der pII B cell epitope peptide.
XX XX
XX KW T cell epitope; B cell epitope; allergy; allergen; antigenic;
XX KW antiallergic; antiasthmatic; antiinflammatory; dermatological;
XX KW immunosuppressive; vaccine; rhinitis; sinusitis; bronchial asthma;
XX KW atopic dermatitis; acute urticaria; chronic urticaria;
XX KW gastro-intestinal syndrome; food allergen; oro-pharyngeal syndrome;
XX KW anaphylactic reaction; drug hypersensitivity; allergic reaction.
XX XX
XX OS Dermatophagoides pteronyssinus.
XX OS Clostridium tetani.

```

```

OS Synthetic.
XX WO200006694-A2.
XX
XX 10-FEB-2000.
XX
XX 20-JUL-1999; 99WO-BE000092.
XX
XX 30-JUL-1998; 98EP-00870167.
XX
XX (UNIO ) UCB SA.
XX
XX Saint-Remy J, Jacquemin M;
XX
XX WPI; 2000-422470/36.
XX
XX New compound for prevention and treatment of allergies comprises at least
XX one allergenic determinant recognized by a B cell and at least
XX one antigenic determinant which does not trigger T cell activation.
XX
XX Claim 8; Page 35; 50pp; English.
XX
XX The present invention describes a compound (I) for the prevention and/or
XX treatment of allergy. The compound comprises at least one allergen
XX antigenic determinant (i) recognised by a B cell or an antibody secreted
XX by a B cell of a non-atopic individual and at least one antigenic
XX determinant (ii) different from the allergen that triggers T cell
XX activation. (i) has antiallergic, antisthmatic, antiinflammatory,
XX dermatological and immunosuppressive activities, and can be used in a
XX vaccine. (i) may be used in a pharmaceutical or cosmetic medicament to
XX treat and/or prevent allergies or a disease of allergic origin,
XX especially hypersensitivities. These include rhinitis, sinusitis,
XX bronchial asthma, atopic dermatitis, some forms of acute and chronic
XX urticaria, gastro-intestinal syndromes associated with the ingestion of
XX food allergens, oro-pharyngeal syndromes, anaphylactic reactions
XX associated with drug hypersensitivities and/or a mixture of these. The
XX use of (i) in the treatment of allergic conditions avoids the need for
XX drug treatment, which often causes undesirable side-effects. Also, prior
XX art drug therapies alleviate symptoms, but do not influence their causes,
XX however (i) actually combats the cause of an allergic reaction. The
XX present sequence represents a specifically claimed compound peptide
XX sequence from the present invention
XX
XX Sequence 31 AA;
XX
XX Query Match 55.1%; Score 81.5; DB 3; Length 31;
XX Best Local Similarity 61.3%; Pred. No. 3.2e-05;
XX Matches 19; Conservative 0; Mismatches 3; Indels 9; Gaps 1;
XX
XX Qy 1 QYIKANSKFIGITELG-----PSLHWS 22
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db 1 QYIKANSKFIGITELGGHEIKKVLVPGCHG 31
XX
XX RESULT 14
XX AAU11423
XX ID AAU11423 standard; peptide; 33 AA.
XX
XX AC AAU11423;
XX
XX 12-MAR-2002 (first entry)
XX
XX Synthetic immunogen peptide 4.
XX
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
XX Plasmodium falciparum.
XX Mammalia.
XX Synthetic.
XX
OS

```

```

OS Chimeric.
XX Key Location/Qualifiers
XX Peptide 1..20
XX /note= "Malaria CSP protein (378-398 aa)"
XX Peptide 21..24
XX /note= "Spacer peptide"
XX Peptide 25..33
XX /note= "Gonadotropin releasing hormone epitope"
XX Modified-site 33
XX /note= "Amidated glycine or glycineamide"
XX
XX WO200185763-A2.
XX
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014363.
XX
XX 05-MAY-2000; 2000US-0202328P.
XX
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX or its analog.
XX
XX Claim 11; Page 8; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known as
XX luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
XX which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX useful inducing an immune response against GnRH in an animal subject, and
XX as such is useful as a contraceptive and in the treatment of diseases
XX such as cancer (of the breast, uterus and other gynaecological cancer),
XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX prostate cancer. The immunogen is effective in eliciting high and
XX specific anti-GnRH antibody titres. The present sequence is a synthetic
XX immunogen of the invention
XX
XX Sequence 33 AA;
XX
XX Query Match 55.1%; Score 81.5; DB 5; Length 33;
XX Best Local Similarity 66.7%; Pred. No. 3.4e-05;
XX Matches 16; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
XX
XX Qy 4 KANSKFIGITELGPSLHWSYGLRP 27
XX ||||| : ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX Db 10 KASSVF-NVANSGPSLHWSYGLRP 32
XX
XX RESULT 15
XX AAO30489
XX ID AAO30489 standard; protein; 194 AA.
XX
XX AC AAO30489;
XX
XX 22-SEP-2003 (first entry)
XX
XX Human TNFalpha variant, TNF34-P30-P2.
XX
XX Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;
XX tumour necrosis factor alpha; gene therapy; arthritis; human; mutant;
XX mutein; variant; tetanus toxoid; epitope.
XX
XX Homo sapiens.
XX Unidentified.
XX
OS

```

```
OS Chimeric.
XX Key Location/Qualifiers
FH Region 1. .109
FT Region /note= "Human TNF"
FT Region 110. .130
FT Region /note= "Tetanus toxoid P30 epitope"
FT Region 131. .145
FT Region /note= "Tetanus toxoid P2 epitope"
FT Region 146. .194
FT Region /note= "Human TNF"
XX WO2003042244-A2.
PN
XX
XX 22-MAY-2003.
XX
XX 15-NOV-2002; 2002WO-DK000764.
XX
XX 16-NOV-2001; 2001DK-00001702.
XX 16-NOV-2001; 2001US-0331575P.
XX
XX (PHAR-) PHARMEXA AS.
XX (KLYS/) KLYSNER S.
XX (NIEL/) NIELSEN F S.
XX (BRAT/) BRATT T.
XX (VOLD/) VOLDORGB B.
XX (MOUR/) MOURITSEN S.
XX
XX Klysner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;
XX WPI; 2003-449558/42.
XX
XX PT New immunogenic analogue of a polymeric protein, useful for preparing a
XX composition for treating inflammatory diseases e.g. arthritis.
XX Claim 23; Page 159-160; 196pp; English.
XX
XX The invention relates to immunogenic analogues of multimeric proteins
XX such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis
XX factor alpha (TNF, TNFalpha) and methods for production of immunogenic
XX analogues. The immunogenic analogue is useful for preparing a composition
XX for treating inflammatory diseases, e.g., arthritis. It is also used in
XX gene therapy. The present sequence is human TNFalpha variant protein with
XX an inserted tetanus toxoid P2 and P30 epitopes. This sequence is used to
XX illustrate the method of the invention
XX
XX Sequence 194 AA;
XX
XX Query Match 54.7%; Score 81; DB 6; Length 194;
XX Best Local Similarity 76.2%; Pred. NO. 0.0003;
XX Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 QYIKANSKFIGITELQPSLHW 21
XX Db 131 QYIKANSKFIGITELAEAKPW 151
XX
XX Search completed: March 10, 2004, 09:12:10
XX Job time : 42.1634 secs
```

GenCore version 5.1.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 27.6732 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QVIKANSKFIGITELGSLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|--------|--------------------|
| 1 | 74 | 50.0 | 1310 | 2 | Q93N27 | Q93n27 clostridium |
| 2 | 53 | 35.8 | 751 | 16 | Q88EP1 | Q88ef1 pseudomonas |
| 3 | 52.5 | 35.5 | 342 | 16 | Q88FZ0 | Q88fz0 rhizobium l |
| 4 | 52 | 35.1 | 91 | 13 | Q8PRH0 | Q8prh0 argulla ja |
| 5 | 52 | 35.1 | 226 | 12 | Q8BAZ8 | Q8baz8 hepatitis b |
| 6 | 52 | 35.1 | 382 | 5 | Q95VY3 | Q95vy3 lumbricus t |
| 7 | 51.5 | 34.8 | 834 | 16 | Q8U7P1 | Q8u7p1 agrobacteri |
| 8 | 51 | 34.5 | 384 | 5 | O77072 | O77072 eisenia foe |
| 9 | 51 | 34.5 | 388 | 17 | Q9YD14 | Q9ydl4 aeropyrum p |
| 10 | 51 | 34.5 | 524 | 4 | Q88Z6 | Q88z6 homo sapien |
| 11 | 51 | 34.5 | 539 | 4 | Q81YK5 | Q81yk5 homo sapien |
| 12 | 50.5 | 34.1 | 60 | 7 | Q31585 | Q31585 salmo salar |
| 13 | 50.5 | 34.1 | 71 | 7 | Q9XBJ9 | Q9xj9 salvelinus |
| 14 | 50.5 | 34.1 | 85 | 7 | Q95IS2 | Q95is2 salmo salar |
| 15 | 50.5 | 34.1 | 85 | 7 | Q95HY1 | Q95hy1 salmo salar |
| 16 | 50.5 | 34.1 | 86 | 7 | Q95HX4 | Q95hx4 salmo salar |

| | | | | | | |
|----|------|------|------|----|--------|--------------------|
| 17 | 50.5 | 34.1 | 244 | 7 | Q31590 | Q31590 salmo salar |
| 18 | 50 | 33.8 | 94 | 13 | Q8UUK6 | Q8uuk6 scleropages |
| 19 | 50 | 33.8 | 274 | 3 | O14264 | O14264 schizosacch |
| 20 | 50 | 33.8 | 344 | 16 | Q89RX8 | Q89rx8 bradyrhizob |
| 21 | 50 | 33.8 | 480 | 16 | Q91482 | Q91482 pseudomonas |
| 22 | 49.5 | 33.4 | 67 | 7 | Q31578 | Q31578 salmo salar |
| 23 | 49 | 33.1 | 103 | 12 | Q8QSV9 | Q8qsv9 hepatitis b |
| 24 | 49 | 33.1 | 324 | 12 | Q89914 | Q89914 sonchus vel |
| 25 | 49 | 33.1 | 484 | 16 | Q8HZ30 | Q8hz30 pseudomonas |
| 26 | 48 | 32.4 | 226 | 12 | Q71672 | Q71672 hepatitis b |
| 27 | 48 | 32.4 | 232 | 16 | Q888R4 | Q888r4 pseudomonas |
| 28 | 48 | 32.4 | 318 | 16 | Q8UE59 | Q8ue59 agrobacteri |
| 29 | 48 | 32.4 | 334 | 16 | Q89Y42 | Q89y42 bradyrhizob |
| 30 | 48 | 32.4 | 794 | 16 | Q82DY5 | Q82dy5 versinia pe |
| 31 | 48 | 32.4 | 1036 | 16 | Q8PIX2 | Q8pix2 xanthomonas |
| 32 | 48 | 32.4 | 2091 | 3 | P78616 | P78616 emericella |
| 33 | 47.5 | 32.1 | 84 | 13 | Q9DEK4 | Q9dek4 coregonus s |
| 34 | 47.5 | 32.1 | 85 | 7 | Q95IS3 | Q95is3 salmo salar |
| 35 | 47.5 | 32.1 | 85 | 7 | Q95IR2 | Q95ir2 salmo salar |
| 36 | 47.5 | 32.1 | 149 | 7 | Q31495 | Q31495 oncorhynch |
| 37 | 47.5 | 32.1 | 205 | 17 | Q981D4 | Q981d4 sulfolobus |
| 38 | 47.5 | 32.1 | 216 | 7 | Q9GJH0 | Q9gjh0 salmo trutt |
| 39 | 47.5 | 32.1 | 216 | 7 | Q9GJG9 | Q9gjj9 salmo trutt |
| 40 | 47.5 | 32.1 | 1361 | 12 | Q8BF56 | Q8bf56 tulip virus |
| 41 | 47 | 31.8 | 114 | 10 | Q8L8X1 | Q8l8x1 arabidopsis |
| 42 | 47 | 31.8 | 134 | 10 | Q9LSN9 | Q9lsn9 arabidopsis |
| 43 | 47 | 31.8 | 230 | 8 | Q8HMC7 | Q8hmc7 physcillus |
| 44 | 47 | 31.8 | 323 | 10 | Q7Y0Z6 | Q7y0z6 lycopersico |
| 45 | 47 | 31.8 | 342 | 16 | Q82TX0 | Q82tx0 nitrosomona |

ALIGNMENTS

RESULT 1
Q93N27 PRELIMINARY; PRT; 1310 AA.
ID Q93N27
AC Q93N27
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RA Shumin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389424; AK72964.2; ...
DR GO; GO:0004866; Ferredoxin-like protein; IEA.
DR GO; GO:0008237; F-metalloproteinase activity; IEA.
DR GO; GO:0015070; F-toxin activity; IEA.
DR GO; GO:0008270; F-zinc ion binding; IEA.
DR GO; GO:0009405; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA-like lec_gl.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR000395; Peptidase_M27.
DR InterPro; IPR006035; Pept M_zn_BS.
DR Pfam; PF01742; Peptidase_M27; I.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PRODOM; PD001963; BONTOTOXILYSIN; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
FT NON TER 1 1310
FT NON TER 1310 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418E450 CRC64;
Query Match 50.0%; Score 74; DB 2; Length 1310;

Best Local Similarity 100.0%; Pred. No. 0.023; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QYKANSKFIGITEL 15
 |||||
 Db 831 QYKANSKFIGITEL 845

RESULT 2

ID Q88EF1 PRELIMINARY; PRT; 751 AA.
 AC Q88EF1;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Conserved hypothetical protein.
 GN PP4514.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=150488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;

RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapfel E., Scanlan D., Tran K., Moazzes A.,
 RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eissen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AE016790; AAN70088.1; -;
 DR TIGR; PP4514; -;
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0045735; P:nutrient reservoir activity; IEA.
 DR InterPro; IPR002641; Patatin.
 DR Pfam; PF01734; Patatin; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 751 AA; 83174 MW; CD140194E4CA6368 CRC64;

Query Match 35.8%; Score 53; DB 16; Length 751;
 Best Local Similarity 58.3%; Pred. No. 23;
 Matches 14; Conservative 4; Mismatches 4; Indels 2; Gaps 2;

Qy 2 YKANSKFIGI-TLGP-SLHWSYG 24

Db 712 YINAASIFGLDTPGP-LNPSYG 734

RESULT 3

ID Q98FZ0 PRELIMINARY; PRT; 342 AA.
 AC Q98FZ0;
 DT 01-OCT-2001 (TRENBLrel. 18, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein mll3560.
 GN MLL3560.
 OS Mesorhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003002; BAB50426.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 342 AA; 37582 MW; 99597D57D25D11A1 CRC64;

Query Match 35.5%; Score 52.5; DB 16; Length 342;
 Best Local Similarity 37.5%; Pred. No. 11;
 Matches 9; Conservative 7; Mismatches 5; Indels 3; Gaps 1;

Qy 1 QYKANSKFIGITELGSLHWSYG 24

Db 82 RFLKAGSDFIGVADTG--YWFPG 102

RESULT 4

ID Q9PRH0 PRELIMINARY; PRT; 91 AA.
 AC Q9PRH0;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Prepro-mGnRH precursor (Gonadoliberin) (Gonadotropin-releasing
 hormone) (LH-RH) (Luliberin).
 OS Anguilla japonica (Japanese eel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
 OC Anguilla.
 OX NCBI_TaxID=7937;
 RN [1]_TaxID=7937;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Okubo K., Suetake H., Aida K.;
 RT "Expression of two gonadotropin-releasing hormone (GnRH) precursor
 RT genes in various tissues of the Japanese eel and evolution of GnRH.";
 RL Zool. Sci. 16:471-478(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Okubo K., Suetake H., Aida K.;
 RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
 RT hormone (prepro-mGnRH) mRNA is present in the brain and various
 RT peripheral tissues of the Japanese eel.";
 RL Zool. Sci. 16:645-651(1999).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GnRH FAMILY.
 DR EMBL; AB026989; BAB82608.1; -;
 DR EMBL; AB026991; BAB83597.1; -;
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadoliberin1.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERN1.
 DR PROSITE; PS00473; GnRH; 1.
 KW Amidation; Hormone; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 32 MGNRH.
 FT CHAIN 33 91 GnRH ASSOCIATED PEPTIDE.
 SQ SEQUENCE 91 AA; 9893 MW; BA15C9DC08434A7B CRC64;

Query Match 35.1%; Score 52; DB 13; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 HWSYGLRP 27

Db 24 HWSYGLRP 31

RESULT 5
Q8BAZ8 PRELIMINARY; PRT; 226 AA.
ID Q8BAZ8;
AC Q8BAZ8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface antigen.
OS Hepatitis B virus.
OC Viruses; Retroviral viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F078;
RA Bowden R.J.;
RT "Hepatitis B virus variability and human population history in the Pacific";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=F078;
RA Basuni A.A.; Bowden R.J.; Butterworth L.-A.; Cookesley G.; Locarnini S.; Carman W.F.;
RL EMBL; AX124488; AAM82819.1; -;
DR GO; GO:0016032; Pivotal life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surfac.
DR Pfam; PF00695; VMSA; 1.
SQ SEQUENCE 226 AA; 25414 MW; 63791C8AEFFB514 CRC64;

Query Match 35.1%; Score 52; DB 12; Length 226;
Best Local Similarity 39.1%; Pred. No. 8.2;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 5 ANSKFIGITELGSLHWSYGLRP 27
DB 166 ASVRFSLWSSLAFFVHMFVGLSP 188
RESULT 6
Q95VY3 PRELIMINARY; PRT; 382 AA.
ID Q95VY3;
AC Q95VY3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coelomic cytolytic factor precursor.
OS Lumbriacus terrestris (Common earthworm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbriquina; Lumbricidae; Lumbricus.
OX NCBI_TaxID=6398;
RN [1]
RP SEQUENCE FROM N.A.
RA Beschlin A.; De Baetselier P.; Bilej M.;
RT "Distinct carbohydrate recognition domains of an earthworm defense molecule recognize Gram negative and Gram positive bacteria";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395805; AAL09587.1; -;
DR GO; GO:0004553; Phyllosilicate activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P-carbohydrate metabolism; IEA.
DR InterPro; IPR008985; ConA like lec.gl.
DR Pfam; PF00722; Glyco_hydro_16; 1.
DR Signal.
FT SIGNAL
SQ SEQUENCE 382 AA; 43931 MW; 5256CF171EB7D3FB CRC64;

Query Match 35.1%; Score 52; DB 5; Length 382;
Best Local Similarity 36.8%; Pred. No. 15;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 6 NSKFIGITELGSLHWSYGLRP 24
DB 196 SGEFIGIQMGSTHMGPG 214
RESULT 7
Q8U7P1 PRELIMINARY; PRT; 834 AA.
ID Q8U7P1;
AC Q8U7P1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Periplasmic nitrate reductase large subunit.
GN NAPA OR ATU4408 OR AGR L 917.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W.; Setubal J.C.; Kaul R.; Monks D.E.; Kitajima J.P.;
RA Chua V.K.; Zhou Y.; Chen L.; Wood G.E.; Almeida N.F. Jr.; Woo L.;
RA Kren V.; Paulsen I.T.; Eisen J.A.; Karp P.D.; Bovee D. Sr.;
RA Chapman P.; Clendenning J.; Deatherage G.; Gillet W.; Grant C.;
RA Kutyavin I.; Levy R.; Li M.-J.; McClelland E.; Palmeri A.; Gordon D.;
RA Raymond C.; Rouse G.; Saenphimmachak C.; Wu Z.; Romero P.; Gordon D.;
RA Zhang S.; Yoo H.; Tao Y.; Biddle P.; Jung M.; Krespan W.; Perry M.;
RA Gordon-Kamm B.; Liao L.; Kim S.; Hendrick C.; Zhao Z.-Y.; Dolan M.;
RA Chumley F.; Tingey S.V.; Tomb J.-F.; Gordon M.P.; Olson M.V.;
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58";
RL Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B.; Hinkle G.; Gattung S.; Miller N.; Blanchard M.;
RA Quorillo B.; Goldman B.S.; Cao Y.; Askenazi M.; Halling C.; Mullin L.;
RA Roumel K.; Gordon J.; Vaudin M.; Lartchouk O.; Epp A.; Liu F.;
RA Wollam C.; Allinger M.; Doughty D.; Scott C.; Lappas C.; Markelz B.;
RA Flanagan C.; Crowell C.; Gursun J.; Lomo C.; Sear C.; Strub G.;
RA Cielo C.; Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58";
RL Science 294:2323-2328 (2001).
DR EMBL; AB009369; AAL45202.1; ALT_INIT.
DR EMBL; AB008245; AAK89031.1; -;
DR PIR; AD3098; AD3098.
DR PIR; E98188; E98188.
DR GO; GO:0006118; P-electron transport; IEA.
DR InterPro; IPR009010; Asp decarb. fold.
DR InterPro; IPR006656; Molybdopterin.
DR InterPro; IPR006965; Molybdopterin.
DR InterPro; IPR006657; Mol dinuc. bind.
DR InterPro; IPR006655; Prox. Mboxred.
DR InterPro; IPR006311; Tat.
DR Pfam; PF00384; molybdopterin; 1.
DR Pfam; PF04879; Molybdopterin; 1.
DR Pfam; PF01568; Molybdopterin; 1.
DR TIGRfam; TIGR01409; TAT signal seq; 1.
DR PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
DR Complete proteome.
SQ SEQUENCE 834 AA; 93527 MW; FF707BC71CA08B9E CRC64;

Query Match 34.8%; Score 51.5; DB 16; Length 834;
Best Local Similarity 37.0%; Pred. No. 44;
Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

QY 2 YKANSKFI-GITELGSLHWSYGLRP 27
DB 289 FVRNHTKFGVGTDIG-----YGLRP 309

```

RESULT 8
O77072 ID O77072 PRELIMINARY; PRT; 384 AA.
AC O77072;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coelomic cytolitic factor 1.
GN CCFL.
OS Eiseia foetida (Common branding worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
OX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98406152; PubMed=9733802;
RA Beschin A., Bilej M., Hanssens F., Raymakers J., Van Dyck E.,
RA Revets H., Brys L., Gomez J., De Baetselier P., Timmermans M.;
RT "Identification and cloning of a glucan- and Lipopolysaccharide-
RT binding protein from Eisenia foetida earthworm involved in the
RT activation of prophenoloxidase cascade.";
RL J. Biol. Chem. 273:24948-24954 (1998).
DR EMBL; AF030028; AAC35887.1; -.
DR GO; GO:0004553; F:Hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:Carbohydrate metabolism; IEA.
DR InterPro; IPR008985; ConsA-like lec gl.
DR InterPro; IPR000757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
DR SEQUENCE 384 AA; 44322 MW; C9085C94003BAD6D CRC64;

Query Match 34.5%; Score 51; DB 5; Length 384;
Best Local Similarity 41.2%; Pred. No. 32;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 8 KFGITELGPSLHWSYG 24
DB 198 EPLGIQKMGSTWHNPG 214

RESULT 9
O9YD14 ID O9YD14 PRELIMINARY; PRT; 388 AA.
AC O9YD14;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 388AA long hypothetical FMU protein.
GN APEI098.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101 (1999).
DR EMBL; AP000060; BAA80083.1; -.
DR PIR; C72710; C72710.
DR GO; GO:0003723; P:RNA binding; IEA.
DR InterPro; IPR002478; PUA.
DR InterPro; IPR001678; Sun_Nop1/Nop2.
DR InterPro; IPR004521; Unchar_dom 2.
DR Pfam; PF01189; Noll_Nop2_Sun; 1.

Pfam; PF01472; PUA; 1.
SMART; SM00359; PUA; 1.
TIGRFAM; TIGR00451; unchar_dom_2; 1.
PROSITE; PS00890; PUA; 1.
KW Complete proteome.
SQ SEQUENCE 388 AA; 42535 MW; A9E10CEAAEF0B0AC CRC64;

Query Match 34.5%; Score 51; DB 17; Length 388;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 14 ELGPSLHWSYGLRP 27
DB 345 ELGRLTWSWGLRP 358

RESULT 10
Q8N826 ID Q8N826 PRELIMINARY; PRT; 524 AA.
AC Q8N826;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ38654.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project."
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
CC DR EMBL; AK095973; BAC04663.1; -.
DR GO; GO:0007155; P:Cell adhesion; IEA.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000421; FA58_C.
DR InterPro; IPR008979; Gal_Bind-like.
DR InterPro; IPR004043; LCCL_dom.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00754; F5_F8-type_C; 1.
DR SMART; SM00231; FA58C; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00603; LCCL; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00022; FA58C_3; 1.
DR PROSITE; PS00820; LCCL; 1.
KW Hypothetical protein.
SQ SEQUENCE 524 AA; 57698 MW; CEAC4F78A492EB44 CRC64;

Query Match 34.5%; Score 51; DB 4; Length 524;
Best Local Similarity 40.9%; Pred. No. 31;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 3 IKANSKFIGITELGPSLHWSYG 24
DB 267 IRASSSQVNSGSDQVHSPG 288

RESULT 11
Q81YK5 ID Q81YK5 PRELIMINARY; PRT; 539 AA.
AC Q81YK5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Similar to RIKEN cDNA 4631413K11 gene.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC035671; AAH35671.1; -
 DR Genew; HGNC:21479; DBELDI.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR008559; CUB.
 DR InterPro; IPR000421; FAS8.C.
 DR InterPro; IPR008979; Gal_bind_like.
 DR InterPro; IPR004043; LCCL_dom.
 DR Pfam; PF00431; CUB; 1.
 DR Pfam; PF00754; F5_P8_type_C; 1.
 DR Pfam; PF03815; LCCL; 1.
 DR SMART; SM00042; CUB; 1.
 DR SMART; SM00231; FAS8C; 1.
 DR SMART; SM00603; LCCL; 1.
 DR PROSITE; PS01180; CUB; 1.
 DR PROSITE; PS00022; FAS8C_3; 1.
 DR PROSITE; PS00820; LCCL; 1.
 SQ SEQUENCE 539 AA; 59196 MW; 929D4BIACIAAAA4C CRC64;

 Query Match 34.5%; Score 51; DB 4; Length 539;
 Best Local Similarity 40.9%; Pred. No. 32;
 Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

 QY 3 IKANSKFIGITELGSLHWSYG 24
 Db 267 IRASSWQSVNESGDQVHWSPG 288

 RESULT 12
 Q31585
 ID Q31585 PRELIMINARY; PRT; 60 AA.
 AC Q31585;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE (DB02) MHC class II beta 1 (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Gimholt U., Oleaker I., de Vries Lindstrom C., Lie O.;
 RA "A study of polymorphism in the MHC class II beta 1 and MHC class I
 RT alpha 2 domain exons of Atlantic salmon (Salmo salar).";
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L24953; AAA49597.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0045012; F:MHC class II receptor activity; IEA.
 DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.
 DR GO; GO:0019886; P:antigen processing, exogenous antigen via M...; IEA.
 DR GO; GO:0006955; P:immune response; IEA.
 DR InterPro; IPR000353; MHC_II_beta.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR ProDom; PD000328; MHC_II_beta; 1.
 KW Glycoprotein; MHC II; Transmembrane.
 FT NON_TER 1
 FT NON_TER 60
 SQ SEQUENCE 60 AA; 6776 MW; 5DAF3449060940E2 CRC64;

 Query Match 34.1%; Score 50.5; DB 7; Length 60;
 Best Local Similarity 57.9%; Pred. No. 3.1;

Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

 QY 1 QYIKANS---KFIGITELG 16
 Db 16 EYIRFNSTVGKFGVGYTELG 34

 RESULT 13
 Q9XRJ9
 ID Q9XRJ9 PRELIMINARY; PRT; 71 AA.
 AC Q9XRJ9;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE MHC class II beta 1 (Fragment).
 GN SANA.
 OS Salvelinus namaycush (lake trout).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
 OX NCBI_TaxID=8040;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Dorschner M.O., Duris T., Phillips R.B.;
 RA "Diversity of a Lake Trout Mhc class II Gene";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF130026; AAD20889.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0045012; F:MHC class II receptor activity; IEA.
 DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.
 DR GO; GO:0019886; P:antigen processing, exogenous antigen via M...; IEA.
 DR GO; GO:0006955; P:immune response; IEA.
 DR InterPro; IPR000353; MHC_II_beta.
 DR Pfam; PF00969; MHC_II_beta; 1.
 DR ProDom; PD000328; MHC_II_beta; 1.
 KW Glycoprotein; MHC II; Transmembrane.
 FT NON_TER 1
 FT NON_TER 71
 SQ SEQUENCE 71 AA; 8002 MW; E3095286582A9F2D CRC64;

 Query Match 34.1%; Score 50.5; DB 7; Length 71;
 Best Local Similarity 57.9%; Pred. No. 3.7;
 Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

 QY 1 QYIKANS---KFIGITELG 16
 Db 14 EYIRFNSTVGKFGVGYTELG 32

 RESULT 14
 Q95IS2
 ID Q95IS2 PRELIMINARY; PRT; 85 AA.
 AC Q95IS2;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE MHC class II beta chain (Fragment).
 OS Salmo salar (Atlantic salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Landry C., Bernatchez L.;
 RA "Comparative analysis of population structure across environments and
 RT geographic scales at Major Histocompatibility Complex and
 RT microsatellite in Atlantic salmon (Salmo salar).";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF373699; AAK51882.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0045012; F:MHC class II receptor activity; IEA.
 DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 8.71595 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-11

Perfect score: 148

Sequence: 1 QYIKANSKFIGITELGPSLHWSYGLRPX 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

- 1: Piri:*
- 2: Piri2:*
- 3: Piri3:*
- 4: Piri4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 74 | 50.0 | 1315 | 1 BTCLTN | tentoxilysin (EC 3 |
| 2 | 54 | 36.5 | 67 | 2 I78541 | gonadoliberin prec |
| 3 | 54 | 36.5 | 90 | 1 RHMSG | gonadoliberin prec |
| 4 | 54 | 36.5 | 92 | 1 RHUG | gonadoliberin prec |
| 5 | 54 | 36.5 | 92 | 1 RHRTG | gonadoliberin prec |
| 6 | 52 | 35.1 | 10 | 1 RHPOG | gonadoliberin - pi |
| 7 | 52 | 35.1 | 10 | 1 RHSGG | gonadoliberin - sh |
| 8 | 52 | 35.1 | 89 | 2 I51423 | gonadoliberin prec |
| 9 | 51.5 | 34.8 | 824 | 2 AD3098 | periplasmic nitrat |
| 10 | 51.5 | 34.8 | 834 | 2 E98188 | probable fmu prote |
| 11 | 51 | 34.5 | 388 | 2 C72710 | class II histocomp |
| 12 | 50.5 | 34.1 | 244 | 2 S29582 | probable MPS trans |
| 13 | 50 | 33.8 | 274 | 2 T39087 | hypothetical prote |
| 14 | 50 | 33.8 | 480 | 2 A83487 | hypothetical prote |
| 15 | 49 | 33.1 | 422 | 2 T38962 | potassium uptake p |
| 16 | 49 | 33.1 | 484 | 2 E83245 | hemocytin - silkw |
| 17 | 49 | 33.1 | 3133 | 2 S52093 | gonadoliberin I - |
| 18 | 48 | 32.4 | 10 | 1 RHAQ1 | gonadoliberin I pr |
| 19 | 48 | 32.4 | 92 | 2 I50644 | transketolase [imp |
| 20 | 48 | 32.4 | 318 | 2 AD2810 | probable transketo |
| 21 | 48 | 32.4 | 318 | 2 G97588 | pyruvate, water di |
| 22 | 48 | 32.4 | 794 | 2 AB0294 | conserved hypothet |
| 23 | 47.5 | 32.1 | 205 | 2 C90140 | hypothetical prote |
| 24 | 47 | 31.8 | 728 | 2 E83228 | hypothetical prote |
| 25 | 46.5 | 31.4 | 193 | 2 F89967 | class II histocomp |
| 26 | 46.5 | 31.4 | 245 | 2 S29980 | hypothetical prote |
| 27 | 46.5 | 31.4 | 388 | 2 A82445 | gonadotropin-relea |
| 28 | 46 | 31.1 | 98 | 2 I50739 | hypothetical prote |
| 29 | 46 | 31.1 | 171 | 2 S38237 | |

adenylate cyclase, glucosamine-fructo C4-dicarboxylate t pyruvate, water di phosphoenolpyruvat protein F14J36.10 probable amidohydr hypothetical prote peptidyl-prolyl ci FKBP-type peptidyl hypothetical prote probable transamin conserved hypothet hypothetical prote hypothetical prote alpha-glucuronidas

30 46 31.1 186 2 A90167
31 46 31.1 349 2 E75611
32 46 31.1 459 2 G82431
33 46 31.1 522 2 T43369
34 46 31.1 780 2 D75361
35 46 31.1 4056 2 H96599
36 45.5 30.7 322 2 T38399
37 45 30.4 256 2 F50616
38 45 30.4 258 2 F72052
39 45 30.4 258 2 B86573
40 45 30.4 368 2 T27432
41 45 30.4 397 2 B70815
42 45 30.4 435 2 C89857
43 45 30.4 456 2 S55661
44 45 30.4 644 2 S46746
45 45 30.4 674 2 H72423

ALIGNMENTS

RESULT 1

BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N/Alternate names: tetanus neurotoxin
C/Species: Clostridium tetani
C/Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text change 03-Jun-2002
C/Accession: A25689; A25757; A25194; B25194; A60759; S63348; S63344
R/Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, EMBO J. 5, 2495-2502, 1986
A/Title: Tetanus toxin: primary structure, expression in E. coli, and homology with bc
A/Reference number: A25689; MUID:87053814; PMID:3536478
A/Accession: A25689
A/Molecule type: DNA
A/Residues: 1-1315 <EIS>
A/Cross-references: GB:K04436; NID:G40769; PIDN:CAA28033.1; PID:G40770
R/Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A/Title: The complete nucleotide sequence of tetanus toxin.
A/Reference number: A25757; MUID:87040747; PMID:3774547
A/Accession: A25757
A/Molecule type: DNA
A/Residues: 1-1315 <FAI>
A/Cross-references: GB:X06214; NID:G40773; PIDN:CAA29564.1; PID:G40774
R/Experimental source: strain CN3911
R/Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A/Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in
A/Reference number: A25194; MUID:86085672; PMID:3510187
A/Accession: A25194
A/Molecule type: DNA
A/Residues: 743-1315 <FA2>
A/Cross-references: GB:M12739; NID:G144920; PIDN:AAA23282.1; PID:G144921
A/Accession: B25194
A/Molecule type: protein
A/Residues: 865-894 <PA3>
R/Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A/Title: Isolation, purification, and characterization of fragment B, the NH-2-termina
A/Reference number: A60759; MUID:90035436; PMID:2478476
A/Accession: A60759
A/Molecule type: Protein
A/Residues: 461-475 <NA>
R/Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A/Title: Delineation of several DR-restricted tetanus toxin T cell epitopes.
A/Reference number: J50098; MUID:89093918; PMID:2463305
R/Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.
Nature 359, 832-835, 1992
A/Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteol
A/Reference number: S27125; MUID:93063293; PMID:1331807
A/Contents: annotation

R/de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Eur. J. Biochem. 229, 61-69, 1995
 A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
 A;Reference number: S69348; MUID:95262688; PMID:7744050
 A;Accession: S69348
 A;Molecule type: protein
 A;Residues: 2-31 <DEF>
 C;Comment: The source of this protein was an extrachromosomal plasmid.
 C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
 dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
 C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to ganglio
 C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
 presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
 C;Function:
 A;Description: blocks neuroexcitotoxicity via hydrolysis of a Gln-Phe peptide bond in synap
 A;Superfamily: tetanus toxin
 C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F;2-457/Product: tetroxylisin light chain (fragment A) #status predicted <TTL>
 F;461-1115/Product: tetroxylisin heavy chain (fragment B) #status experimental <TTH>
 F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
 F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
 F;233,237/Binding site: zinc (His) #status predicted
 F;234/Active site: Glu #status predicted

Query Match 50.0%; Score 74; DB 1; Length 1315;

Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYIKANSKFIGITEL 15
 DB 830 QYIKANSKFIGITEL 844

RESULT 2

178541
 Gonadoliberin precursor - rhesus macaque (fragment)
 N;Alternate names: luteinizing hormone releasing hormone
 C;Species: Macaca mulatta (rhesus macaque)
 C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
 A;Accession: I78541
 R;Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
 Neuroendocrinology 60, 346-359, 1994
 A;Title: Developmental expression of the genes encoding transforming growth factor alpha
 A;Reference number: 158134; MUID:95124501; PMID:7545971
 A;Accession: I78541
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-67 <RES>
 A;Cross-references: GB:S75918; NID:G912831; PIDN:AAB33096.1; PID:G912832
 C;Superfamily: gonadoliberin

Query Match 36.5%; Score 54; DB 2; Length 67;
 Best Local Similarity 90.0%; Pred. No. 0.69;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
 DB 5 SQHWSYGLRP 14

RESULT 3

RHWSG
 Gonadoliberin precursor - mouse
 N;Alternate names: gonadotropin-releasing hormone (GNRH); luteinizing hormone releasing
 C;Species: Mus musculus (house mouse)
 C;Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
 A;Accession: A47578
 R;Ma, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolic,
 Science 234, 1366-1371, 1986
 A;Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo
 A;Reference number: A47578; MUID:87069928; PMID:3024317
 A;Accession: A47578

A;Molecule type: DNA
 A;Residues: 1-90 <MAS>
 A;Cross-references: EMBL:ML4872; NID:G1935576; PIDN:AAA37717.1; PID:G387175
 C;Genetics:
 A;Introns: 45/3; 77/3
 C;Function:
 A;Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropir
 A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
 C;Superfamily: gonadoliberin
 C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;22-31/Product: gonadoliberin #status predicted <GLB>
 F;35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
 F;22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte
 F;31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gl

Query Match 36.5%; Score 54; DB 1; Length 90;

Best Local Similarity 90.0%; Pred. No. 0.95;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
 DB 21 SQHWSYGLRP 30

RESULT 4

RHUG
 gonadoliberin precursor [validated] - human
 N;Alternate names: gonadotropin releasing hormone (GNRH); luteinizing hormone releasing
 C;Species: Homo sapiens (man)
 C;Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 08-Dec-2000
 A;Accession: S05308; A26173; A93342; A90108; A01410; S45718
 R;Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
 Nucleic Acids Res. 17, 6403-6404, 1989
 A;Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone g
 A;Reference number: S05308; MUID:89366682; PMID:2671939
 A;Accession: S05308
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-92 <HAY>
 A;Cross-references: EMBL:X15215; NID:G91955; PIDN:CAA33285.1; PID:G91956
 R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonad
 A;Reference number: A94090; MUID:86094338; PMID:2867548
 A;Accession: A26173
 A;Molecule type: mRNA
 A;Residues: 1-92 <ADE>
 A;Cross-references: GB:M12578; NID:G183418; PIDN:AAA35916.1; PID:G386749
 A;Experimental source: hypothalamus
 R;Seeburg, P.H.; Adelman, J.P.
 Nature 311, 666-668, 1984
 A;Title: Characterization of cDNA for precursor of human luteinizing hormone releasing
 A;Reference number: A93342; MUID:85012739; PMID:6090951

A;Accession: A93342

A;Molecule type: mRNA

A;Residues: 1-15, 'S', 17-92 <SEE>

A;Cross-references: GB:X01059; NID:G34356; PIDN:CAA25526.1; PID:G34357

A;Experimental source: placenta

R;Tan, L.; Rousseau, P.

Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982

A;Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized i

A;Reference number: A90108; MUID:83126573; PMID:6760865

A;Accession: A90108

A;Molecule type: protein

A;Residues: 24-33 <FAN>

A;Experimental source: placental trophoblasts

R;Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterc

FEBS Lett. 346, 203-206, 1994

A;Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by t

A;Reference number: S45718; MUID:94283597; PMID:8013634

A;Accession: A47578

A;Contents: annotation; degradation pathway of synthetic hormone

C:Genetics:

A:Gene: GDB:GNRH; LHRH: GRH
 A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
 A:Map position: 8p21-8p11.2
 A:Introns: 47/3; 79/3
 C:Function:
 A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
 A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-92/Product: progadoliberin #status predicted <PGN>
 F:24-33/Product: gonadoliberin #status experimental <NAT>
 F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
 F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
 F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 36.5%; Score 54; DB 1; Length 92;

Best Local Similarity 90.0%; Pred. No. 0.97; 1; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27

DB 23 SQHWSYGLRP 32

RESULT 5

RHRTG

Gonadoliberin precursor - rat
 N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormone
 N:Contains: gonadoliberin; prolactin release-inhibiting factor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
 C:Accession: A40147; B26173; A48410
 R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
 Mol. Endocrinol. 3, 1257-1262, 1989
 A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
 A:Reference number: A40147; MUID:89384561; PMID:2476669
 A:Accession: A40147
 A:Molecule type: DNA
 A:Residues: 1-92 <BON>
 A:Cross-references: GB:M31670; NID:G204447; PIDN:AAA41264.1; PID:G204448
 R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
 Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
 A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
 A:Reference number: A94090; MUID:86094338; PMID:2867548
 A:Accession: B26173
 A:Molecule type: mRNA
 A:Residues: 1-92 <ADE>
 A:Cross-references: GB:M12579; NID:G204445; PIDN:AAA41263.1; PID:G204446
 R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
 Cell. Mol. Neurobiol. 12, 447-454, 1992
 A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
 A:Reference number: A48410; MUID:93105480; PMID:1468115
 A:Accession: A48410
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-92 <MAI>
 A:Cross-references: GB:S50870; NID:G262059; PIDN:AB24572.1; PID:G262060
 A:Experimental source: thymus
 A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)
 C:Genetics:
 A:Introns: 47/3; 79/3
 C:Function:

A:Description: stimulates pituitary secretion of lutropin and follitropin
 A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid; r
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-92/Product: progadoliberin #status predicted <PGN>
 F:24-33/Product: gonadoliberin #status predicted <GAP>
 F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
 F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted <PIF>

F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following g
 Query Match 36.5%; Score 54; DB 1; Length 92;
 Best Local Similarity 90.0%; Pred. No. 0.97;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27

DB 23 SQHWSYGLRP 32

RESULT 6

RHFGG

Gonadoliberin - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
 C:Accession: A01411
 R:Baba, Y.; Matsuo, H.; Schally, A.V.
 Biochem. Biophys. Res. Commun. 44, 459-463, 1971
 A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of t
 A:Reference number: A90172; MUID:72114303; PMID:4946067
 A:Accession: A01411
 A:Molecule type: protein
 A:Residues: 1-10 <BAB>
 R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
 Biochem. Biophys. Res. Commun. 45, 822-827, 1971
 A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase met
 A:Reference number: A90176; MUID:72065376; PMID:4942726
 A:Contents: annotation; synthesis
 A:Note: the synthetic and natural hormones have the same physicochemical and biologica
 R:Baba, Y.; Arimura, A.; Schally, A.V.
 Biochem. Biophys. Res. Commun. 45, 483-487, 1971
 A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
 A:Reference number: A90175; MUID:72117544; PMID:4946275
 A:Contents: annotation
 A:Note: Trp-3 appears to be essential for biological activity
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and

QY 20 HWSYGLRP 27

DB 2 HWSYGLRP 9

Query Match 35.1%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7

RHSHG

Gonadoliberin - sheep
 C:Species: Ovis orientalis aries. Ovis ammon aries (domestic sheep)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
 C:Accession: A93780; A01411
 R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
 Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
 A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fac
 A:Reference number: A93780; MUID:72094314; PMID:4550508
 A:Accession: A93780
 A:Molecule type: protein
 A:Residues: 1-10 <BUR>

A:Note: the natural and synthetic hormones have the same biological activity
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 35.1%; Score 52; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
 |||||
 Db 2 HWSYGLRP 9

RESULT 8
 151423
 gonadoliberin precursor - African clawed frog
 N:Alternate names: luteinizing hormone releasing hormone
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
 C:Accession: 151423
 R:Hayes, W.P.; Wray, S.; Battey, J.F.
 Endocrinology 134, 1835-1845, 1994
 A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain
 A:Reference number: 151423; MUID:94185563; PMID:8137750
 A:Accession: 151423
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-89 <HAY>
 A:Cross-references: GB:I28040; NID:G496291; PIDN:AAA49728.1; PID:G496292
 C:Genetics:
 A:Gene: GnRH-I
 C:Superfamily: gonadoliberin

Query Match 35.1%; Score 52; DB 2; Length 89;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
 |||||
 Db 25 HWSYGLRP 32

RESULT 9
 AD3098
 periplasmic nitrate reductase large subunit [imported] - Agrobacterium tumefaciens (stra
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AD3098
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 i; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AD3098
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-824 <KUR>
 A:Cross-references: GB:AB2577; MUID:21608550; PMID:11743193
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: napA
 A:Map position: linear chromosome
 C:Superfamily: Alcaligenes eutrophus nitrate reductase A chain

Query Match 34.8%; Score 51.5; DB 2; Length 824;
 Best Local Similarity 37.0%; Pred. No. 24;
 Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

QY 2 YIKANSKFI-GITELGPSLHWSYGLRP 27
 ...:|||||
 Db 279 FVRNHTKFRGVTDIG-----YGLRP 299

RESULT 10
 E98188
 periplasmic nitrate reductase precursor napA (AF040988) [imported] - Agrobacterium tumef

C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C:Accession: E98188
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldmar
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: E98188
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-834 <KUR>
 A:Cross-references: GB:AB007870; PIDN:AAK89031.1; PID:G15158825; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR L 917
 A:Map position: linear chromosome
 C:Superfamily: Alcaligenes eutrophus nitrate reductase A chain

Query Match 34.8%; Score 51.5; DB 2; Length 834;
 Best Local Similarity 37.0%; Pred. No. 24;
 Matches 10; Conservative 9; Mismatches 1; Indels 7; Gaps 2;

QY 2 YIKANSKFI-GITELGPSLHWSYGLRP 27
 ...:|||||
 Db 289 FVRNHTKFRGVTDIG-----YGLRP 309

RESULT 11
 C72710
 probable fmu protein APE1098 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C:Accession: C72710
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: C72710
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-388 <KAW>
 A:Cross-references: DBJ:AP000060; NID:G5104188; PIDN:BAA80083.1; PID:G1043869; PID:G51
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1098

Query Match 34.5%; Score 51; DB 2; Length 388;
 Best Local Similarity 71.4%; Pred. No. 12;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 14 ELGPSLHWSYGLRP 27
 |||||
 Db 345 ELGPSLHWSYGLRP 358

RESULT 12
 S29982
 class II histocompatibility antigen - Atlantic salmon
 C:Species: Salmo salar (Atlantic salmon)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
 C:Accession: S29982
 R:Hardvik, I.
 submitted to the EMBL Data Library, October 1992
 A:Reference number: S29980
 A:Accession: S29982
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-244 <HOR>
 A:Cross-references: EMBL:X70166; NID:G64369; PID:G64370
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 34.1%; Score 50.5; DB 2; Length 244;

```

Best Local Similarity 57.9%; Pred. No. 8.9;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 QYIKANS---KFIGITELG 16
    :|||: |||: |||: |||:
Db 51 EYIRENSTVGKFKVGVTELG 69

RESULT 13
T39087
hypothetical protein SPAC7D4.09c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39087
R:Genies, S.; Churcher, C.M.; Wood, V.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21826
A:Accession: T39087
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <GEN>
A:Cross-references: EMBL:Z99532; PIDN:CAB16726.2; GSPDB:GN00066; SPDB:SPAC7D4.09c
C:Genetics:
A:Gene: SPDB:SPAC7D4.09c
A:Map position: 1

Query Match 33.8%; Score 50; DB 2; Length 274;
Best Local Similarity 29.4%; Pred. No. 12;
Matches 10; Conservative 7; Mismatches 9; Indels 8; Gaps 1;

QY 2 YIKANSKFIGITELGPSLHW-----SYGLRP 27
    :|||: |||: |||: |||:
Db 18 YFTSTVTLFVSLKNAFSLSLWLMKYGGHONFGLKP 51

RESULT 14
A83487
probable MFS transporter PA1262 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83487
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950, MUID:20437337; PMID:10984043
A:Accession: A83487
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <STO>
A:Cross-references: GB:AE004556; GB:AE004091; NID:ig9947194; PIDN:AAG04651.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1262
C:Superfamily: lincomycin-resistance protein lmrB

Query Match 33.8%; Score 50; DB 2; Length 480;
Best Local Similarity 88.9%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 LGPSLHWSY 23
    |||||: |||:
Db 224 LGPSLHWSW 232

RESULT 15
T38962
hypothetical protein SPAC5D6.04 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38962

R:Skelton, J.; Churcher, C.M.; Wood, V.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1997
A:Reference number: Z21820
A:Accession: T38962
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-452 <SKE>
A:Cross-references: EMBL:Z98056; PIDN:CAB10852.1; GSPDB:GN00066; SPDB:SPAC5D6.04
A:Experimental source: strain 972h-; cosmid cSD6
C:Genetics:
A:Gene: SPDB:SPAC5D6.04
A:Map position: 1
A:Introns: 48/3; 92/2; 122/3

Query Match 33.1%; Score 49; DB 2; Length 452;
Best Local Similarity 57.1%; Pred. No. 29;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 13 TELGPSLHWSYGLR 26
    :|||: |||: |||: |||:
Db 166 SQLGQALRWSYGYR 179

Search completed: March 10, 2004, 09:16:47
Job time : 9.77478 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 5.12062 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-11
Perfect score: 148
Sequence: 1 QYIKANSKFIGITELGSLHWSYGLRPX 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 74 | 50.0 | 1314 | 1 | TEXX_CLOTE |
| 2 | 54 | 36.5 | 67 | 1 | GONI_MACMO |
| 3 | 54 | 36.5 | 90 | 1 | GONI_MOUSE |
| 4 | 54 | 36.5 | 91 | 1 | GONI_PIG |
| 5 | 54 | 36.5 | 92 | 1 | GONI_HUMAN |
| 6 | 54 | 36.5 | 92 | 1 | GONI_RAT |
| 7 | 52 | 35.1 | 61 | 1 | GONI_SHEEP |
| 8 | 52 | 35.1 | 63 | 1 | GONI_MESAU |
| 9 | 52 | 35.1 | 89 | 1 | GONI_XENLA |
| 10 | 52 | 35.1 | 90 | 1 | GONI_RANCA |
| 11 | 52 | 35.1 | 92 | 1 | GONI_TUPGB |
| 12 | 49 | 33.1 | 452 | 1 | YDQ4_SCHPO |
| 13 | 49 | 33.1 | 3133 | 1 | HMCT_BOMMO |
| 14 | 48 | 32.4 | 10 | 1 | GONI_ALIMI |
| 15 | 48 | 32.4 | 92 | 1 | GONI_CAVPO |
| 16 | 48 | 32.4 | 92 | 1 | GONI_CHICK |
| 17 | 47 | 31.8 | 216 | 1 | YDQ4_LEPIN |
| 18 | 46 | 31.1 | 94 | 1 | GONI_HAPBU |
| 19 | 46 | 31.1 | 95 | 1 | GONI_MORSA |
| 20 | 46 | 31.1 | 95 | 1 | GONI_PAGMA |
| 21 | 46 | 31.1 | 95 | 1 | GONI_SPAAU |
| 22 | 46 | 31.1 | 99 | 1 | YHBU_ACTAC |
| 23 | 46 | 31.1 | 110 | 1 | YHBU_DICLA |
| 24 | 46 | 31.1 | 459 | 1 | DCUC_VIBCH |
| 25 | 46 | 31.1 | 780 | 1 | PPSA_DEIRA |
| 26 | 45.5 | 30.7 | 322 | 1 | YABU_SCHPO |
| 27 | 45 | 30.4 | 258 | 1 | MIP_CHLPN |
| 28 | 45 | 30.4 | 347 | 1 | A11C_MOUSE |
| 29 | 45 | 30.4 | 644 | 1 | YHJ9_YEAST |
| 30 | 45 | 30.4 | 674 | 1 | AGUA_THEMEA |
| 31 | 45 | 30.4 | 831 | 1 | NAPA_ALCEU |
| 32 | 44 | 29.7 | 66 | 1 | VG84_BPMLS |
| 33 | 44 | 29.7 | 90 | 1 | GON8_RANDY |

ALIGNMENTS

RESULT 1

| ID | TEXX_CLOTE | STANDARD; | PRT; | 1314 AA. |
|----|--|-----------------------------------|------|----------|
| AC | P04958; | 1987 (Rel. 05, Created) | | |
| DT | 13-AUG-1987 | (Rel. 05, Last sequence update) | | |
| DT | 13-AUG-1987 | (Rel. 42, Last annotation update) | | |
| DT | 10-OCT-2003 | (Rel. 42, Last annotation update) | | |
| DE | Tetanus toxin precursor [EC 3.4.24.68] (Tentoxylisin) [Contains: | | | |
| DE | Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy | | | |
| DE | chain (Tetanus toxin chain H)]. | | | |
| GN | TEXX OR CTP60. | | | |
| OS | Clostridium tetani. | | | |
| OC | Plasmid pE88, and Plasmid 75 Kbp. | | | |
| OC | Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; | | | |
| OC | Clostridium. | | | |
| OX | NCBI_TaxID=1513; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | PLASMID=75 Kbp; | | | |
| RC | MEDLINE=87053814; PubMed=3536478; | | | |
| RA | Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J., | | | |
| RA | Weller U., Hudel M., Habermann E., Niemann H.; | | | |
| RT | "Tetanus toxin: primary structure, expression in E. coli, and | | | |
| RT | homology with botulinum toxins."; | | | |
| RL | EMBO J. 5:2495-2502(1986). | | | |
| RL | Nucleic Acids Res. 14:7809-7812(1986). | | | |
| RA | Fairweather N.F., Lyness V.A.; | | | |
| RA | MEDLINE=87040747; PubMed=3774547; | | | |
| RA | "The complete nucleotide sequence of tetanus toxin."; | | | |
| RA | [3] | | | |
| RA | SEQUENCE FROM N.A. | | | |
| RA | STRAIN=CN3911; PLASMID=75 Kbp; | | | |
| RA | MEDLINE=22457253; PubMed=12552129; | | | |
| RA | Brueggemann H., Baumer S., Fricke W.F., Wierse A., Liesegang H., | | | |
| RA | Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A., | | | |
| RA | Gottschalk G.; | | | |
| RA | "The genome sequence of Clostridium tetani, the causative agent of | | | |
| RA | tetanus disease."; | | | |
| RA | Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003). | | | |
| RA | [4] | | | |
| RA | SEQUENCE OF 742-1314 FROM N.A. | | | |
| RA | PLASMID=75 Kbp; | | | |
| RA | MEDLINE=86085672; PubMed=3510187; | | | |
| RA | Fairweather N.F., Lyness V.A., Fickard D.J., Allen G., Thomson R.O.; | | | |
| RA | "Cloning, nucleotide sequencing, and expression of tetanus toxin | | | |
| RA | fragment C in Escherichia coli."; | | | |
| RA | J. Bacteriol. 165:21-27(1986). | | | |
| RA | [5] | | | |
| RA | PARTIAL SEQUENCE, AND DISULFIDE BONDS. | | | |
| RA | MEDLINE=30201034; PubMed=2108021; | | | |
| RA | Kriegstein K., Henschen A., Weller U., Habermann E.; | | | |
| RA | "Arrangement of disulfide bridges and positions of sulfhydryl groups | | | |
| RA | in tetanus toxin."; | | | |

P37545 bacillus su
Q58778 methanococ
O78450 guillardia
Q09305 caenorhabdi
P53083 saccharomyc
P19721 mouse adeno
P56161 anopheles s
P32176 escherichia
Q91409 dicentrarch
Q9kct3 bacillus ha
Q06737 bacillus su
Q82267 chlamydomphi

RL Eur. J. Biochem. 188:39-45(1990).
 RN [6]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=92037649; PubMed=1935979;
 RA Krieglstein K.G., Henschen A.H., Weller U., Habermann E.;
 RT "limited proteolysis of tetanus toxin. Relation to activity and
 identification of cleavage sites.";
 RL Eur. J. Biochem. 202:41-51(1991).
 RN [7]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010949; PubMed=1396558;
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.;
 RT "tetanus toxin is a zinc protein and its inhibition of
 neurotransmitter release and protease activity depend on zinc.";
 RL EMBO J. 11:3577-3583(1992).
 RN [8]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "tetanus and botulinum-B neurotoxins block neurotransmitter release
 by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 neurotoxin.";
 RL Nat. Struct. Biol. 4:788-792(1997).
 CC -!- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 BOND OF SYNAPTOSOMAL-2.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
 synaptobrevin 2.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 AND ARE NON-TOXIC AFTER SEPARATION.
 CC -!- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 GLYCOSIDE RECEPTORS.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04436; CAA28033.1; -;
 DR EMBL; X06214; CAA29564.1; -;
 DR EMBL; AF528097; AA037454.1; -;
 DR EMBL; M12739; AA23282.1; -;
 DR FIR; A25689; BTCLTN.
 DR PDB; IAF9; 29-APR-98.
 DR PDB; IASD; 14-OCT-98.
 DR PDB; IDOH; 27-MAR-00.
 DR PDB; IDFQ; 24-MAR-00.
 DR PDB; IDIW; 24-MAR-00.
 DR PDB; IDLL; 24-MAR-00.
 DR PDB; IFV3; 05-SEP-01.
 DR MEROPS; M27.001; -;
 DR InterPro; IPR008985; Consa_like lec_gl.
 DR InterPro; IPR002160; Kunitz legume.
 DR InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR000395; Peptidase_M27.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; P000760; BONTOKILYSIN.
 DR PRODOM; P0001963; Bontokilysin; 1.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT-MET 0
 FT CHAIN 1 456 TETANUS TOXIN LIGHT CHAIN.
 FT CHAIN 457 1314 TETANUS TOXIN HEAVY CHAIN.
 FT METAL 232 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 236 ZINC (CATALYTIC) (BY SIMILARITY).
 FT TRANSMEM 226 246 POTENTIAL.
 FT TRANSMEM 669 689 POTENTIAL.
 FT DISULFID 438 466 INTERCHAIN.
 FT DISULFID 1076 1092
 FT HELIX 876 882
 FT TURN 883 883
 FT STRAND 884 891
 FT TURN 892 893
 FT STRAND 894 897
 FT STRAND 904 907
 FT TURN 909 910
 FT STRAND 912 915
 FT STRAND 920 925
 FT TURN 928 929
 FT STRAND 932 935
 FT HELIX 938 940
 FT TURN 941 946
 FT STRAND 949 956
 FT HELIX 962 968
 FT TURN 969 970
 FT STRAND 972 977
 FT STRAND 980 981
 FT HELIX 983 985
 FT STRAND 987 995
 FT TURN 996 997
 FT STRAND 998 1004
 FT TURN 1006 1007
 FT STRAND 1010 1016
 FT STRAND 1020 1020
 FT TURN 1021 1022
 FT STRAND 1031 1037
 FT TURN 1039 1040
 FT STRAND 1042 1047
 FT TURN 1048 1049
 FT STRAND 1050 1056
 FT TURN 1058 1059
 FT STRAND 1068 1074
 FT TURN 1079 1080
 FT STRAND 1082 1091
 FT HELIX 1097 1105
 FT TURN 1106 1107
 FT STRAND 1112 1112
 FT STRAND 1114 1114
 FT TURN 1116 1117
 FT STRAND 1120 1120
 FT STRAND 1122 1122
 FT TURN 1123 1124
 FT STRAND 1127 1131
 FT HELIX 1132 1134
 FT TURN 1135 1136
 FT STRAND 1137 1141
 FT TURN 1144 1145
 FT STRAND 1148 1152
 FT STRAND 1155 1158
 FT TURN 1159 1162
 FT STRAND 1163 1166
 FT TURN 1173 1178
 FT STRAND 1184 1185
 FT STRAND 1188 1188
 FT STRAND 1190 1190

Query Match 50.0%; Score 74; DB 1; Length 1314;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYKANSKFIGITEL 15
DB 829 QYKANSKFIGITEL 843

RESULT 2

ID GON1_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)]
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I] (Fragment).
DE GNRH1 OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth factor alpha and its receptor in the hypothalamus of female rhesus macaques";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both luteinizing and follicle-stimulating hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; S75918; AAB33096.1; -
DR PIR; I78541; I78541.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 1
FT NON_TER 1 1
FT CHAIN 6 >67
FT PEPTIDE 6 15
FT ACT_SITE 19 >67
FT MOD_RES 6 6
FT MOD_RES 15 15
FT MOD_RES 67 67
FT NON_TER 67 67
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 36.5%; Score 54; DB 1; Length 67;
Best Local Similarity 90.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
DB 5 SCHWSYGLRP 14

RESULT 3

ID GON1_MOUSE STANDARD; PRT; 90 AA.
AC F13562;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)]
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor I].
DE GNRH1 OR GNRH.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Haylick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolics K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse";
RL Science 234:1366-1371(1986).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both luteinizing and follicle-stimulating hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M14872; AAA37717.1; -
DR PIR; A47578; RMSG.
DR MGD; MGI:95789; GnRH.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadolibereinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERNI.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 21
FT CHAIN 22 90
FT NON_TER 22 90
FT PEPTIDE 22 31
FT ACT_SITE 35 90
FT MOD_RES 24 24
FT MOD_RES 22 22
FT MOD_RES 31 31
FT MOD_RES 90 90
SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 36.5%; Score 54; DB 1; Length 90;
Best Local Similarity 90.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27
DB 5 SCHWSYGLRP 14

Db 21 SQHWSYGLRP 30

RESULT 4

GN1_PIG STANDARD; PRT; 91 AA.

AC P4921; 1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Progonadolibarin I precursor [Contains: Gonadolibarin I (LH-RH I)]

DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I].

GN GN1 OR GN1H

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Hypothalamus;

RA Weesner G.D., Matteri R.L., Becker B.A.;

RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 24-33.

RX MEDLINE=72117544; PubMed=4946067;

RA Baba Y., Matsuo H., Schally A.V.;

RT "Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the proposed structure by conventional sequential analyses".

RL Biochem. Biophys. Res. Commun. 44:459-463 (1971).

RN [3]

RP SYNTHESIS OF GONADOLIBERIN.

RX MEDLINE=72065376; PubMed=4942726;

RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;

RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase method".

RL Biochem. Biophys. Res. Commun. 45:822-827 (1971).

RN [4]

RP SYNTHESIS OF GONADOLIBERIN.

RX MEDLINE=72117544; PubMed=4946275;

RA Baba Y., Arimura A., Schally A.V.;

RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";

RL Biochem. Biophys. Res. Commun. 45:483-487 (1971).

CC !- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both luteinizing and follicle-stimulating hormones.

CC !- SUBCELLULAR LOCATION: Secreted.

CC !- SIMILARITY: Belongs to the GnRH family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; L32864; AAA31066.1; .

DR InterPro; IPR002012; GnRH.

DR Pfam; PF004079; GonadolibarinI.

DR PRINTS; PR01541; GnRH; 1.

DR PROSITE; PS00473; GnRH; 1.

KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Signal; Pyrrolidone carboxylic acid.

FT SIGNAL 1 23

FT CHAIN 24 91

FT PEPTIDE 24 33

FT ACT_SITE 26 26

FT MOD_RES 24 24

FT MOD_RES 33 33

SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

AMIDATION (G-34 PROVIDE AMIDE GROUP).

Query Match 36.5%; Score 54; DB 1; Length 91;

Best Local Similarity 90.0%; Pred. No. 0.32;

Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 18 SLHWSYGLRP 27

DB 23 SQHWSYGLRP 32

RESULT 5

GN1_HUMAN STANDARD; PRT; 92 AA.

AC P01138;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Progonadolibarin I precursor [Contains: Gonadolibarin I (LH-RH I)]

DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated peptide I].

GN GN1 OR GN1H OR LHRH.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89366682; PubMed=2671939;

RA Hayflick J.S., Adelman J.P., Seeburg P.H.;

RT "The complete nucleotide sequence of the human gonadotropin-releasing hormone gene.";

RL Nucleic Acids Res. 17:6403-6403 (1989).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=86094338; PubMed=2867548;

RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;

RT "Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat.";

RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183 (1986).

RN [3]

RP SEQUENCE FROM N.A. AND VARIANT SER-16.

RX MEDLINE=85012739; PubMed=6090951;

RA Seeburg P.H., Adelman J.P.;

RT "Characterization of cDNA for precursor of human luteinizing hormone releasing hormone.";

RL Nature 311:666-668 (1984).

RN [4]

RP SEQUENCE OF 24-33.

RX MEDLINE=83126573; PubMed=6760865;

RA Tan L., Rousseau P.;

RT "The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in the human placenta.";

RL Biochem. Biophys. Res. Commun. 109:1061-1071 (1982).

RN [5]

RP VARIANT SER-16.

RX MEDLINE=99318093; PubMed=10391209;

RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N., Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L., Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q., Lander E.S.;

RT "Characterization of single-nucleotide polymorphisms in coding regions of human genes.";

RL Nat. Genet. 22:231-238 (1999).

RN [6]

RP ERRATUM.

RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N., Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L., Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q., Lander E.S.;

RL Nat. Genet. 23:373-373(1999).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACOLOGICAL: Available under the names Factrel (Ayerst Labs),
 CC Luterpulse or Lutrelf (Perring Pharmaceuticals) and Relisorm
 CC (Serono).
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X1059; CAA5526.1; -;
 DR EMBL; M12578; AAA35916.1; -;
 DR EMBL; X15215; CAA33285.1; -;
 DR PIR; S05308; RHUG.
 DR Genew; HGNC:4419; GNRH1.
 DR MIM; 152760; -;
 DR GO; GO:0005625; C:soluble fraction; TAS.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadolibnerin.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBNERI.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pharmaceutical; Signal; Polymorphism;
 KW Pyrrolidone carboxylic acid.
 FT SIGNAL 1 23
 FT CHAIN 24 92 PROGNADOLIBNERIN I.
 FT PEPTIDE 24 33 GONADOLIBNERIN I.
 FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACT_SITE 26 26 ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 FT VARIANT 16 16 W -> S (in dbSNP:6185).
 FT FTID=VAR 013943.
 SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
 Query Match 36.5%; Score 54; DB 1; Length 92;
 Best Local Similarity 90.0%; Pred. No. 0.32;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 18 SLHWSYGLRP 27
 Db 23 SQHWSYGLRP 32
 RESULT 6
 CONL_RAT STANDARD; PRT; 92 AA.
 AC P07490;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Prognadolibnerin I precursor [Contains: Gonadolibnerin I (LH-RH I)
 DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 DE hormone I) (GnRH I) (Luliberin I); Prolactin release-inhibiting factor
 DE I].
 GN GNRH1 OR GNRH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RT Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9384661; PubMed=2476669;
 RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and
 RT hypothalamic expression";
 RT Mol. Endocrinol. 3:1257-1262(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Thymus;
 MEDLINE=93105480; PubMed=1468115;
 RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
 RT "Thymocytes express a mRNA that is identical to hypothalamic
 RT luteinizing hormone-releasing hormone mRNA.";
 RL Cell. Mol. Neurobiol. 12:447-454(1992).
 RN [4]
 RP SEQUENCE OF 1-47 FROM N.A.
 RX TISSUE=Heart;
 MEDLINE=87149087; PubMed=3547652;
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
 RT "Two mammalian genes transcribed from opposite strands of the same
 RT DNA locus.";
 RL Science 235:1514-1517(1987).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Central nervous system.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S08870; AAB24572.1; -;
 DR EMBL; M12579; AAA41263.1; -;
 DR EMBL; M31670; AAA41264.1; -;
 DR EMBL; M15527; AAA42141.1; ALT_SEQ.
 DR EMBL; M15529; AAA42139.1; -;
 DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
 DR PIR; A40147; RHETG.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadolibnerin.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBNERI.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 23
 FT CHAIN 24 92 PROGNADOLIBNERIN I.
 FT PEPTIDE 24 33 GONADOLIBNERIN I.
 FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACT_SITE 26 26 ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
 Query Match 36.5%; Score 54; DB 1; Length 92;
 Query Match


```

RESULT 9
GONI_XENLA
ID GONI_XENLA STANDARD; PRT; 89 AA.
AC P45566
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L28040; AAA49728.1; -
CC PIR; I51423; I51423.
CC InterPro; IPR002012; GNRH.
CC InterPro; IPR004079; GonadoliberinI.
CC Pfam; PF00446; GNRH; 1.
CC PRINTS; PR01541; GONADOLIBRNI.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW SIGNAL; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 89
FT PEPTIDE 24 33
FT PEPTIDE 37 89
FT PEPTIDE 37 85
FT PEPTIDE 24 24
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
SQ SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;

Query Match 35.1%; Score 52; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
DB 25 HWSYGLRP 32

RESULT 10
GONI_RANCA
ID GONI_RANCA STANDARD; PRT; 90 AA.
AC Q90Y63;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LHRH I)
DE Luteinizing hormone-releasing hormone I] (Gonadotropin-releasing

```

```

DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (luliberin I); GNRH-associated peptide I (GAP1)).
GN GNRH1 OR GNRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE=Forebrain;
RX MEDLINE=21102951; PubMed=11170016;
RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GNRH1 and GNRH2
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:190-201(2001).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Forebrain.
CC -!- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
CC during post-breeding. Not expressed in pituitary.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF188754; AAL05972.1; -
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
CC GO; GO:0003755; P:hormone mediated signaling; NAS.
CC GO; GO:0000003; P:reproduction; NAS.
CC InterPro; IPR002012; GNRH.
CC InterPro; IPR004079; GonadoliberinI.
CC Pfam; PF00446; GNRH; 1.
CC PRINTS; PR01541; GONADOLIBRNI.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Signal;
KW Pyroglutamate carboxylic acid.
FT SIGNAL 1 24
FT CHAIN 25 90
FT PEPTIDE 25 34
FT PEPTIDE 38 86
FT MOD_RES 25 25
FT MOD_RES 34 34
FT MOD_RES 34 34
FT MOD_RES 34 34
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;

Query Match 35.1%; Score 52; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 HWSYGLRP 27
DB 26 HWSYGLRP 33

RESULT 11
GONI_TUPGB
ID GONI_TUPGB STANDARD; PRT; 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE Luteinizing hormone-releasing hormone I] (Gonadotropin-releasing

```

```

DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide II.
GN GNRH1 OR GNRH.
OS Tupaia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=37347;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus.
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal.";
CC Gen. Comp. Endocrinol. 104:7-19 (1996).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U63326; AAB16837.1; -.
CC InterPro; IPR002012; GNRH.
CC InterPro; IPR004079; Gonadoliberini.
CC Pfam; PF00446; GNRH; 1.
CC PRINTS; PR01541; GONADOLIBERNI.
CC PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Pictorial; Signal; Pyrrolidone carboxylic acid.
CC SIGNAL 1 23 BY SIMILARITY:
CC FT CHAIN 24 92 GONADOLIBERIN I.
CC FT PEPTIDE 24 33 GONADOLIBERIN I.
CC FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
CC FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
CC FT MOD_RES 33 33 SIMILARITY)
CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
CC FT SIMILARITY).
CC SQ SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
Query Match 35.1%; Score 52; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 HWSYGLRP 27
DB 25 HWSYGLRP 32
RESULT 12
YDQ4_SCHPO STANDARD; PRT; 452 AA.
AC O14197;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C5D6.04 in chromosome I.
GN SPAC5D6.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=972; PubMed=11859360;
RX MEDLINE=21848401; Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Wood V., Gwilliam R., Rajandream M.A., Baker S., Basham D., Bowman S.,
RA Sgourou J., Peat N., Hayles J., Chillingworth T., Churcher C.M.,
RA Brooks K., Brown D., Brown S., Davis P., Feltwell T., Fraser A.,
RA Collins M., Connor R., Cronin A., Harris D., Hidalgo J., Hodgson G.,
RA Grotz S., Goble A., Hamlin N., Harris D., Hunt S., Jagels K.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen O., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO YEAST YER287W.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z98056; CAB10852.1; -.
CC FIR; T38962; T38962.
CC GenesDB_Spombe; SPAC5D6.04; -.
CC InterPro; IPR004776; Auxin eff.
CC Pfam; PF03547; Auxin eff; 1.
CC Hypothetical protein; Transmembrane.
CC FT TRANSMEM 18 38 POTENTIAL.
CC FT TRANSMEM 81 101 POTENTIAL.
CC FT TRANSMEM 269 289 POTENTIAL.
CC FT TRANSMEM 317 337 POTENTIAL.
CC FT TRANSMEM 354 374 POTENTIAL.
CC FT TRANSMEM 390 410 POTENTIAL.
CC FT TRANSMEM 428 448 POTENTIAL.
CC SQ SEQUENCE 452 AA; 49575 MW; 71B77EA5725C69A8 CRC64;
Query Match 33.1%; Score 49; DB 1; Length 452;
Best Local Similarity 57.1%; Pred. No. 9.2;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 13 TELGPSLHWSYGLR 26
DB 166 SQLQALRWSYGYR 179
RESULT 13
HMCT_BOMMO
ID HMCT_BOMMO STANDARD; PRT; 3133 AA.
AC P98032;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hemocytin precursor (Humoral lectin).

```

OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
 OC Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=FUYOU X Tokai; TISSUE=Hemocyte;
 RX MEDLINE=55178544; PubMed=7873598;
 RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
 RA Matsubara F., Tanai K., Kadoh-Okuda K., Kato Y., Mori H.;
 RT "Cloning and expression of the gene of hemocytin, an insect humoral
 RT lectin which is homologous with the mammalian von Willebrand
 RT factor.";
 RL Biochim. Biophys. Acta 1260:245-258(1995).
 RN [2]
 RP SEQUENCE OF 2221-3133 FROM N.A.
 RA Kotani E., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
 RA Matsubara F., Yamakawa M.;
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Adhesive protein and relates to hemostasis or
 CC encapsulation of foreign substances for self-defense.
 CC -1- DEVELOPMENTAL STAGE: Expressed in hemocytes during larVal-pupal
 CC metamorphosis.
 CC -1- INDUCTION: Hemagglutination activity is increased by bacterial
 CC or viral infection and inhibited by D-mannose, N-acetyl-D-
 CC galactosamine and D-maltose.
 CC -1- PTM: May be converted into the 260 kDa mature hemocytin by
 CC proteolysis.
 CC -1- SIMILARITY: TO MAMMALIAN VON WILLEBRAND FACTOR.
 CC -1- SIMILARITY: THIS N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT
 CC OF HUMAN MUCIN 2.
 CC -1- SIMILARITY: Contains 2 F5/8 type C domains.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; D29738; BAA06160.1; -;
 DR ENBL; D14035; BAA03124.1; -;
 DR PIR; S52093; S52093.
 DR HSP; P12259; ICZT.
 DR InterPro; IPR006207; Cys knot C.
 DR InterPro; IPR000421; FASB C.
 DR InterPro; IPR008973; Gal Bind like.
 DR InterPro; IPR002172; LDL receptor_A.
 DR InterPro; IPR002919; TIL Cysrich.
 DR InterPro; IPR001007; VWF C.
 DR InterPro; IPR001846; VWF D.
 DR Pfam; PF00754; F5_F8_type_C; 2.
 DR Pfam; PF01826; TIL; 2.
 DR Pfam; PF00094; vwd; 3.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00231; FAS8C; 2.
 DR SMART; SM00192; LDLA; 1.
 DR SMART; SM00216; VWD; 3.
 DR PROSITE; PS01185; CTCK 1; 1.
 DR PROSITE; PS01225; CTCK 2; 1.
 DR PROSITE; PS01285; FAS8C 1; 2.
 DR PROSITE; PS01285; FAS8C 2; 2.
 DR PROSITE; PS00022; FAS8C 3; 2.
 DR PROSITE; PS01209; VWF 1; FALSE NEG.
 DR PROSITE; PS01209; VWF 1; Signal; Repeat; Cell adhesion.
 KW Lectin; Glycoprotein; Signal; Repeat; Cell adhesion.
 FT SIGNAL ?
 FT CHAIN ? 3133 POTENTIAL.
 FT DOMAIN 29 131 D'.
 FT DOMAIN 153 240 D'.

FT DOMAIN 248 613
 FT DOMAIN 940 1095
 FT DOMAIN 1116 1254
 FT DOMAIN 1283 1356
 FT DOMAIN 1620 1951
 FT DOMAIN 1952 2315
 FT DOMAIN 2320 2321
 FT DOMAIN 2335 2361
 FT DOMAIN 2435 2469
 FT DOMAIN 2553 2622
 FT DOMAIN 2842 2907
 FT DOMAIN 2971 3076
 FT DOMAIN 895 914
 FT DOMAIN 1267 1270
 FT DOMAIN 1425 1428
 FT DOMAIN 1447 1450
 FT DOMAIN 1474 1479
 FT DOMAIN 2148 2153
 FT DOMAIN 2156 2159
 FT DOMAIN 2341 2344
 FT DISULFID 940 1095
 FT DISULFID 1116 1254
 FT DISULFID 2981 3040
 FT DISULFID 2991 3054
 FT DISULFID 3004 3070
 FT DISULFID 3020 3072
 FT DISULFID ? 3075
 FT CARBOHYD 151 151
 FT CARBOHYD 237 237
 FT CARBOHYD 564 564
 FT CARBOHYD 1170 1170
 FT CARBOHYD 1387 1387
 FT CARBOHYD 1622 1622
 FT CARBOHYD 1727 1727
 FT CARBOHYD 1847 1847
 FT CARBOHYD 1975 1975
 FT CARBOHYD 1985 1985
 FT CARBOHYD 2093 2093
 FT CARBOHYD 2113 2113
 FT CARBOHYD 2161 2161
 FT CARBOHYD 2276 2276
 FT CARBOHYD 2451 2451
 FT CARBOHYD 2647 2647
 FT CARBOHYD 2654 2654
 FT CARBOHYD 2663 2663
 FT CARBOHYD 2794 2794
 FT CARBOHYD 2810 2810
 FT CARBOHYD 2865 2865
 FT CARBOHYD 2929 2929
 FT CARBOHYD 2964 2964
 FT CARBOHYD 3028 3028
 FT VARIANT 1288 1288
 FT VARIANT 1305 1305
 SQ SEQUENCE 3133 AA; 343350 MW; E5210D5D14A7B2B2 CRC64;
 Query Match 33.1%; Score 49; DB 1; Length 3133;
 Best Local Similarity 50.0%; Pred. No. 65;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 9 FIGITELGPSLHWSYGLR 26
 DB 346 FLDVPSLGLMSLQWDRGLR 363
 RESULT 14
 GONI ALLMI
 ID GONI ALLMI STANDARD; PRT; 10 AA.
 AC P37041; P20407;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin I (Gonadotropin-releasing hormone I) (GNRH-I) (LH-RH I)
 DE (Luliberin I).

```

DR InterPro; IPR002012; GnrH.
DR InterPro; IPR004079; Gonadolibnerin.
DR PRINTS; PRO1541; GONADOLIBNERI.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 27 32
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT FT MOD_RES 33 33
FT FT SEQUENCE 92 AA; 10279 MW;  ACF74613F456D663 CRC64;
SQ
Query Match 32.4%; Score 48; DB 1; Length 92;
Best Local Similarity 57.1%; Pred. No. 2.6;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 14 ELGPSLHWSYGLRP 27
DB 19 ENGSGQYWSYGVRP 32

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 13.1634 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-10
Perfect score: 186
Sequence: 1 FNNFTVSWLRVPKVSASHLEGSLHWSYGLRPX 34

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/iaa/5A_COMB.pep: *
4: /cgn2_6/ptodata/2/iaa/5B_COMB.pep: *
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep: *
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------------|
| 1 | 152 | 81.7 | 32 | 1 | US-08-446-692-14 |
| 2 | 152 | 81.7 | 32 | 2 | US-08-488-351A-14 |
| 3 | 116 | 62.4 | 188 | 4 | US-09-396-937-14 |
| 4 | 114 | 61.3 | 173 | 4 | US-09-396-937-20 |
| 5 | 112 | 60.2 | 21 | 1 | US-07-610-525-1 |
| 6 | 112 | 60.2 | 21 | 2 | US-08-661-052-12 |
| 7 | 112 | 60.2 | 21 | 2 | US-08-460-502-8 |
| 8 | 112 | 60.2 | 21 | 2 | US-08-724-774B-5 |
| 9 | 112 | 60.2 | 21 | 3 | US-09-089-595-5 |
| 10 | 112 | 60.2 | 21 | 3 | US-09-382-855-5 |
| 11 | 112 | 60.2 | 21 | 3 | US-09-183-714B-5 |
| 12 | 112 | 60.2 | 21 | 3 | US-09-188-082-12 |
| 13 | 112 | 60.2 | 21 | 3 | US-09-171-969-10 |
| 14 | 112 | 60.2 | 21 | 4 | US-09-364-088-12 |
| 15 | 112 | 60.2 | 21 | 4 | US-09-642-281-5 |
| 16 | 112 | 60.2 | 21 | 4 | US-09-102-716-12 |
| 17 | 112 | 60.2 | 21 | 4 | US-08-432-483A-3 |
| 18 | 112 | 60.2 | 21 | 4 | US-09-148-711A-8 |
| 19 | 112 | 60.2 | 21 | 4 | US-09-589-717-5 |
| 20 | 112 | 60.2 | 21 | 4 | US-08-945-289-3 |
| 21 | 112 | 60.2 | 21 | 4 | US-09-396-937-35 |
| 22 | 112 | 60.2 | 21 | 4 | US-09-405-986A-2 |
| 23 | 112 | 60.2 | 21 | 5 | PCT-US93-11703-66 |
| 24 | 112 | 60.2 | 31 | 5 | PCT-US93-11703-64 |
| 25 | 112 | 60.2 | 452 | 1 | US-07-618-312A-2 |
| 26 | 112 | 60.2 | 452 | 1 | US-07-618-312A-4 |
| 27 | 112 | 60.2 | 452 | 1 | US-08-110-786A-8 |

| | | | | | | |
|----|------|------|------|---|-------------------|-------------------|
| 28 | 112 | 60.2 | 452 | 1 | US-08-280-228-2 | Sequence 2, Appl |
| 29 | 112 | 60.2 | 452 | 1 | US-08-280-228-4 | Sequence 4, Appl |
| 30 | 112 | 60.2 | 618 | 1 | US-08-668-381A-5 | Sequence 5, Appl |
| 31 | 112 | 60.2 | 853 | 4 | US-08-913-880C-17 | Sequence 17, Appl |
| 32 | 112 | 60.2 | 858 | 4 | US-08-913-880C-16 | Sequence 16, Appl |
| 33 | 112 | 60.2 | 860 | 4 | US-08-913-880C-15 | Sequence 15, Appl |
| 34 | 112 | 60.2 | 862 | 4 | US-08-913-880C-14 | Sequence 14, Appl |
| 35 | 112 | 60.2 | 865 | 4 | US-08-913-880C-13 | Sequence 13, Appl |
| 36 | 112 | 60.2 | 866 | 4 | US-08-913-880C-12 | Sequence 12, Appl |
| 37 | 112 | 60.2 | 874 | 4 | US-08-913-880C-11 | Sequence 11, Appl |
| 38 | 112 | 60.2 | 875 | 4 | US-08-913-880C-10 | Sequence 10, Appl |
| 39 | 112 | 60.2 | 1315 | 4 | US-08-913-880C-1 | Sequence 1, Appl |
| 40 | 107 | 57.5 | 22 | 1 | US-08-446-692-5 | Sequence 5, Appl |
| 41 | 107 | 57.5 | 22 | 2 | US-08-488-351A-5 | Sequence 5, Appl |
| 42 | 107 | 57.5 | 22 | 3 | US-09-100-409A-41 | Sequence 41, Appl |
| 43 | 107 | 57.5 | 22 | 5 | PCT-US95-13841-8 | Sequence 8, Appl |
| 44 | 100 | 53.8 | 19 | 1 | US-07-610-525-2 | Sequence 2, Appl |
| 45 | 94.5 | 50.8 | 20 | 2 | US-08-319-704-11 | Sequence 11, Appl |

ALIGNMENTS

RESULT 1
US-08-446-692-14
; Sequence 14, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-14

Query Match 81.7%; Score 152; DB 1; Length 32;
Best Local Similarity 87.9%; Pred. No. 1.8e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 FNNFTVSWLRVPKVSASHLEGSLHWSYGLRP 33
Db 3 FNNFTVSWLRVPKVSASHLEGSLHWSYGLRP 31

ADDRESSEE: SHEA & GOULD
STREET: 1251 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES
ZIP: 10020-1193
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/610,525
FILING DATE: 19901108
CLASSIFICATION: 424
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-827-3000
TELEFAX: 212-840-6702
TELEX: 423973
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acid residues
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: no
FRAGMENT TYPE: internal fragment
US-07-610-525-1

Query Match 60.2%; Score 112; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSWLRVPKVSASHLE 21
DB 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 6
US-08-661-052-12
Sequence 12, Application US/08661052
Patent No. 5837243
GENERAL INFORMATION:
APPLICANT: Yashwant M. Deo
APPLICANT: Joel Goldstein
APPLICANT: Robert Graziano
APPLICANT: Chezian Somasundaram
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
OF ANTI-PC RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/661,052
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/484,172
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MXI-043CP

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-661-052-12

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSWLRVPKVSASHLE 21
DB 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 7
US-08-460-502-8
Sequence 8, Application US/08460502
Patent No. 5843464
GENERAL INFORMATION:
APPLICANT: Bakaletz, Lauren O.
APPLICANT: Kaumaya, Parvin T.
TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter and Griswold
STREET: 800 Superior Avenue
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,502
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Golrick, Mary E.
REGISTRATION NUMBER: 34,829
REFERENCE/DOCKET NUMBER: 22727/00120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 822-8458
TELEFAX: (216) 241-0816
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-502-8

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSWLRVPKVSASHLE 21
DB 1 FNNFTVSWLRVPKVSASHLE 21

RESULT 8
US-08-724-774B-5

```
; Sequence 5, Application US/08724774B
; Patent No. 5908778
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
; APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl
; TITLE OF INVENTION: WAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precursors Mage-10,
; TITLE OF INVENTION: Antibodies Specific to The Molecule, and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08724,774B
; FILING DATE: 03-October-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Haneon, No. 5908778man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 688-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-724-774B-5

Query Match 60.2%; Score 112; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 9
; Sequence 5, Application US/09089595
; Patent No. 6153728
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie,
; APPLICANT: Pierre; Cerrottini, Jean-Charles; Carrel,
; APPLICANT: Stefan; Reed, Daryl
; TITLE OF INVENTION: WAGE-10 ENCODING cDNA, The Tumor
; TITLE OF INVENTION: Rejection Antigen Precursors Mage-10,
; TITLE OF INVENTION: Antibodies Specific to The Molecule, and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
```

```
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6153728man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 688-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
; US-09-089-595-5

Query Match 60.2%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FNNFTVSFWLRVPKVSASHLE 21
Db 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 10
; Sequence 5, Application US/09382855
; Patent No. 6174692
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie, Pierre;
; APPLICANT: Cerrottini, Jean-Charles; Carrel, Stefan; Reed, Daryl
; TITLE OF INVENTION: WAGE-10 ENCODING cDNA, The Tumor Rejection
; TITLE OF INVENTION: Antigen Precursors Mage-10, Antibodies Specific To The Molecule
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/382,855
; FILING DATE: 25-August-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/089,595
; FILING DATE: 02-June-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,774
; FILING DATE: 03-October-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Schofield, Mary Anne
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/NAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 21
; TYPE: amino acid
; TOPOLOGY: linear
US-09-382-855-5

Query Match      60.2%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred.No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRPKVKSASHLE 21
Db 1 FNNFTVSFWLRPKVKSASHLE 21

RESULT 11
US-09-183-714B-5
Sequence 5, Application US/09183714B
Patent No. 6221593
GENERAL INFORMATION:
APPLICANT: Boon-Falleur, Thierry
APPLICANT: Brasseur, Francis
APPLICANT: Rimoldi, Donata
APPLICANT: De Plaen, Etienne
TITLE OF INVENTION: Method for Determining Cancer by Determining Expression
TITLE OF INVENTION: of MAGE-10
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/183,714B
CURRENT FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: US 08/724,774
PRIOR FILING DATE: 1996-10-03
NUMBER OF SEQ ID NOS: 7
SEQ ID NO 5
LENGTH: 21
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
US-09-183-714B-5

Query Match      60.2%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred.No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRPKVKSASHLE 21
Db 1 FNNFTVSFWLRPKVKSASHLE 21

RESULT 12
US-09-188-082-12
Sequence 12, Application US/09188082
Patent No. 6270765
GENERAL INFORMATION:
APPLICANT: Yashwant M. Deo
APPLICANT: Joel Goldstein
APPLICANT: Robert Graziano
APPLICANT: Chezian Somsundaram
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
TITLE OF INVENTION: OF ANTI-FC RECEPTOR ANTIBODIES
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,969
FILING DATE: 01 May 1997 (01.05.97)
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/640,713
FILING DATE: 01 May 1996 (01.05.96)
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/802,967
FILING DATE: 21 February 1997 (21.02.97)
ATTORNEY/AGENT INFORMATION:
NAME: Leon R. Yankwich
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: TCS 414.1 PCT (05872)
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL:
ANTI-SENSE:
FEATURE:
NAME/KEY:
APPLICATION NUMBER: US/09/188,082

```

LOCATION:
US-09-171-969-10

Query Match 60.2%; Score 112; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 14

US-09-364-088-12
; Sequence 12, Application US/09364088
; Patent No. 6365161
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo, et al.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street, 24th Floor
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/364,088
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/188,082
; FILING DATE: 07-JUNE-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: MXI-043CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-7414
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-364-088-12

Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
DB 1 FNNFTVSFWLRVPKVSASHLE 21

RESULT 15

US-09-642-281-5
; Sequence 5, Application US/09642281
; Patent No. 6387698
; GENERAL INFORMATION:
; APPLICANT: Rimoldi, Donata; Jongeneel, Victor; Coullie, Pierre;

;; Cerrottini, Jean-Charles; Carrel, Stefan; Reed, Daryl
;; TITLE OF INVENTION: MAGE-10 ENCODING CDNA, The Tumor Rejection
;; Antigen Precursors Mage-10, Antibodies Specific To The Molec
;;
;; NUMBER OF SEQUENCES: 5
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fulbright & Jaworski LLP
;; STREET: 666 Fifth Avenue
;; CITY: New York City
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10103
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
;; COMPUTER: IBM
;; OPERATING SYSTEM: PC-DOS
;; SOFTWARE: Wordperfect
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/642,281
;; FILING DATE: 18-Aug-2000
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/089,595
;; FILING DATE: 02-June-1998
;; APPLICATION NUMBER: US 08/724,774
;; FILING DATE: 03-October-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schofield, Mary Anne
;; REGISTRATION NUMBER: 36,669
;; REFERENCE/DOCKET NUMBER: LUD 5457.2 DIV - JEL/MAS
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 318-3000
;; TELEFAX: (212) 752-5958
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 21
;; TYPE: amino acid
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-642-281-5
Query Match 60.2%; Score 112; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FNNFTVSFWLRVPKVSASHLE 21
|||||
DB 1 FNNFTVSFWLRVPKVSASHLE 21
Search completed: March 10, 2004, 09:28:54
Job time : 14.1634 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:59:54 ; Search time 2.32296 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29

Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/2/iaa/6C_COMB.pep:*
- 6: /cgn2_6/prodata/2/iaa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 29 | 100.0 | 75 | 4 | US-09-621-976-5469 |
| 2 | 29 | 100.0 | 147 | 4 | US-09-252-991A-24435 |
| 3 | 29 | 100.0 | 200 | 3 | US-09-181-183-38 |
| 4 | 29 | 100.0 | 200 | 4 | US-09-280-040-38 |
| 5 | 29 | 100.0 | 200 | 4 | US-09-277-700-38 |
| 6 | 29 | 100.0 | 200 | 4 | US-09-874-585D-38 |
| 7 | 29 | 100.0 | 575 | 4 | US-09-252-991A-17187 |
| 8 | 27 | 93.1 | 690 | 4 | US-09-907-794A-49 |
| 9 | 27 | 93.1 | 690 | 4 | US-09-905-125A-49 |
| 10 | 27 | 93.1 | 890 | 4 | US-09-902-775A-49 |
| 11 | 27 | 93.1 | 888 | 2 | US-08-956-242-4 |
| 12 | 27 | 93.1 | 888 | 3 | US-09-351-215-4 |
| 13 | 26 | 89.7 | 20 | 3 | US-08-658-657B-16 |
| 14 | 26 | 89.7 | 20 | 3 | US-08-763-226C-16 |
| 15 | 26 | 89.7 | 20 | 3 | US-09-307-200-16 |
| 16 | 26 | 89.7 | 20 | 4 | US-09-593-321-16 |
| 17 | 26 | 89.7 | 20 | 4 | US-09-030-619-117 |
| 18 | 26 | 89.7 | 21 | 6 | 5204096-18 |
| 19 | 26 | 89.7 | 76 | 4 | US-09-205-258-892 |
| 20 | 26 | 89.7 | 119 | 4 | US-09-252-991A-24599 |
| 21 | 26 | 89.7 | 179 | 4 | US-09-461-325-345 |
| 22 | 26 | 89.7 | 179 | 4 | US-10-012-542-345 |
| 23 | 26 | 89.7 | 215 | 4 | US-09-252-991A-23779 |
| 24 | 26 | 89.7 | 320 | 4 | US-09-716-129-77 |
| 25 | 26 | 89.7 | 332 | 4 | US-09-198-452A-375 |
| 26 | 26 | 89.7 | 353 | 4 | US-09-716-129-169 |
| 27 | 26 | 89.7 | 410 | 4 | US-09-328-352-5085 |

```

28      26      89.7      477      4      US-09-252-991A-17030      A
29      26      89.7      480      4      US-09-543-681A-7592      Ap
30      26      89.7      802      4      US-09-173-151A-33      Appl
31      26      89.7      855      4      US-08-890-865A-10      Appl
32      26      89.7      971      3      US-09-107-149-19      Appl
33      26      89.7      1036      3      US-08-968-752B-4      Appl
34      26      89.7      1036      4      US-09-536-224-4      Appl
35      26      89.7      1074      2      US-08-768-147B-2      Appl
36      26      89.7      1074      3      US-08-968-752B-2      Appl
37      26      89.7      1074      3      US-09-107-149-3      Appl
38      26      89.7      1074      4      US-09-536-224-2      Appl
39      26      89.7      1663      2      US-08-793-126-1      Appl
40      26      89.7      1663      3      US-09-132-271-1      Appl
41      26      89.7      1663      3      US-09-142-334-22      Appl
42      26      89.7      3594      4      US-09-911-842A-4      Appl
43      26      89.7      3969      3      US-08-661-376-5      Appl
44      25      86.2      14      1      US-07-837-892-1      Appl
45      25      86.2      15      1      US-07-837-892-2      Appl

```

ALIGNMENTS

```

RESULT 1
US-09-621-976-5469
; Sequence 5469, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5469
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -70..-1
US-09-621-976-5469

```

```

Query Match      100.0%; Score 29; DB 4; Length 75;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 SSGPSL 6
Db      41 SSGPSL 46

```

```

RESULT 2
US-09-252-991A-24435
; Sequence 24435, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24435
; LENGTH: 147

```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24435

Query Match      100.0%; Score 29; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSGPSL 6
Db      48 SSGPSL 53

RESULT 3
US-09-181-183-38
; Sequence 38, Application US/09181183
; Patent No. 6146866
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL VEIKKO
; APPLICANT: BACOT, KAREN ONLEY
; TITLE OF INVENTION: LUMAZINE SYNTHASE AND
; RIBOFLAVIN SYNTHASE
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/280.040
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1083
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: M. grisea LS
; APPLICATION NUMBER: US/09/181.183
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1083
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: M. grisea LS
US-09-181-183-38

Query Match      100.0%; Score 29; DB 3; Length 200;
Best Local Similarity 100.0%; Pred. No. 1-2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSGPSL 6
Db      79 SSGPSL 84

RESULT 4
US-09-280-040-38
; Sequence 38, Application US/09280040
; Patent No. 6323013
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL VEIKKO
```

```
; APPLICANT: BACOT, KAREN ONLEY
; APPLICANT: JORDAN, DOUGLAS BRIAN
; TITLE OF INVENTION: LUMAZINE SYNTHASE AND
; RIBOFLAVIN SYNTHASE
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/280.040
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1083
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: M. grisea LS
US-09-280-040-38

Query Match      100.0%; Score 29; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSGPSL 6
Db      79 SSGPSL 84

RESULT 5
US-09-277-700-38
; Sequence 38, Application US/09277700
; Patent No. 6350597
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: BACOT, KAREN O.
; APPLICANT: JORDAN, DOUGLAS B.
; TITLE OF INVENTION: RIBOFLAVIN SYNTHASE GENES AND ENZYMES
; TITLE OF INVENTION: AND METHODS OF USE
; FILE REFERENCE: CL-1083-B
; CURRENT APPLICATION NUMBER: US/09/277,700
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 08/912,218
; EARLIER FILING DATE: AUGUST 15, 1997
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Magnaporthe grisea
US-09-277-700-38

Query Match      100.0%; Score 29; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
```


Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
|||||
Db 79 SSGPSL 84

RESULT 6

US-09-874-585D-38
; Sequence 38, Application US/09874585D
; Patent No. 6682891
; GENERAL INFORMATION:
; APPLICANT: E.I. DuPont de Nemours and Company, Inc.
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: BACOT, KAREN O.
; APPLICANT: JORDAN, DOUGLAS B.
; TITLE OF INVENTION: RIBOFLAVIN SYNTHASE GENES AND ENZYMES AND METHODS OF USE
; FILE REFERENCE: CL-1083-B
; CURRENT APPLICATION NUMBER: US/09/874,585D
; CURRENT FILING DATE: 1997-08-15
; PRIOR APPLICATION NUMBER: US 08/912,218
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Magnaporthe grisea
US-09-874-585D-38

Query Match 100.0%; Score 29; DB 4; Length 200;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
|||||
Db 79 SSGPSL 84

RESULT 7

US-09-252-991A-17187
; Sequence 17187, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17187
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17187

Query Match

Best Local Similarity 100.0%; Score 29; DB 4; Length 575;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
|||||
Db 6 SSGPSL 11

RESULT 8

US-09-907-794A-49
; Sequence 49, Application US/09907794A

; Patent No. 6635458
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10456-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1993-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 49
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-794A-49

```
Query Match          93.1%; Score 27; DB 4; Length 690;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 SSGPSL 6
Db      414 SSGPSI 419

RESULT 9
US-09-905-125A-49
; Sequence 49, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911

Query Match          93.1%; Score 27; DB 4; Length 690;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 SSGPSL 6
Db      414 SSGPSI 419

RESULT 10
US-09-902-775A-49
; Sequence 49, Application US/09902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,775A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
```

```
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 49
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-902-775A-49

Query Match 93.1%; Score 27; DB 4; Length 690;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
Db 414 SSGPSI 419

RESULT 11
US-08-956-242-4
; Sequence 4, Application US/08956242C
; Patent No. 5986081
; GENERAL INFORMATION:
; APPLICANT: Ganetzky, Barry S.
; APPLICANT: Titus, Steven A.
; TITLE OF INVENTION: Polynucleotides Encoding Herg-3
; FILE REFERENCE: 960296.94550
; CURRENT APPLICATION NUMBER: US/08/956,242C
; CURRENT FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (133)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (181)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (438)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (439)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (567)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (568)

; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (571)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (575)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (580)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (589)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (593)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (599)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (600)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (608)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (610)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (611)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (621)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (624)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (625)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (628)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (629)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (634)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (637)
; OTHER INFORMATION: Unidentified at time of filing
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (646)
```

[illegible]